Université catholique de Louvain
École Polytechnique de Louvain

Center for Operations Research and Econometrics

Coordination of Transmission and Distribution System Operations in Electricity Markets

Ilyès Mezghani

Doctoral Thesis

Supervisor:
Anthony Papavasiliou
(UCLouvain, Belgium)

Jury:
Mathieu Van Vyve (UCLouvain, Belgium)
François Glineur (UCLouvain, Belgium)
Burak Kocuk (Sabancı University, Turkey)
Hélène Le Cadre (VITO, Belgium)

President:
Mathieu Van Vyve (UCLouvain, Belgium)

PhD Organization

Ilyès Mezghani
Université catholique de Louvain
École Polytechnique de Louvain
Center for Operations Research and Econometrics

Thesis Supervisors

Anthony Papavasiliou
Associate Professor, Université catholique de Louvain
École Polytechnique de Louvain
Center for Operations Research and Econometrics

Supervisory Committee

Mathieu Van Vyve
Professor, Université catholique de Louvain
Louvain School of Management
Center for Operations Research and Econometrics

François Glineur
Professor, Université catholique de Louvain
École Polytechnique de Louvain
Center for Operations Research and Econometrics
Institute of Information and Communication Technologies, Electronics and Applied Mathematics

Abstract

The integration of renewable energy resources leads to an important change in the way electricity markets are operated and organized. The common approach to the optimization of electric power system operations has focused on the high-voltage transmission network, while the distribution network is typically not accounted for in detail. Nevertheless, renewable resources, like solar panels, are now being integrated at a large scale in the distribution network. The 'distribution' nature of this network is then changing to a broader task: the distribution network still serves power to residential consumers but should also optimize the integration of renewable resources in the system.

Several aspects of this transformation motivate to this dissertation. The contributions are organized in five chapters.

After an introduction of the context in chapter 1, chapter 2 addresses two aspects of interest in power system operations. We first propose a method for solving the alternating current optimal power flow problem, one of the most crucial optimization problems arising in power systems. We rely on Gauss-Newton theory to propose an iterative algorithm, with each iteration solving one or several second-order cone problems. The suggested approach has been extensively tested on MATPOWER instances and reaches competitive performances relative to the state-of-the-art solver IPOPT. We also show the potential improvements that we can attain using warm-start. In the second part of chapter 2, we focus on a stochastic version of the same optimal power flow problem in order to model uncertainty, mainly originating from the renewable resources. We derive a practical and data-driven method for attacking this problem. In particular, detecting critical scenarios can significantly reduce the scenario space. We illustrate this phenomenon on a 1,354-bus system where we reduced the scenario space to 30 scenarios.

In chapter 3, we design five coordination schemes modeling the interactions between transmission and distribution system operators. We specifically analyze the real-time market where the decision process focuses on balancing energy activations and we consider a transmission and distribution network. On a simple example, we analyze the potential strengths and weaknesses of each scheme and we further elaborate on four out of the five schemes in chapters 4 and 5.

In chapter 4, we develop a transmission and distribution market clearing platform. The platform takes into account the complexity of the physics of the networks and non-convex bid structures, and is able to return a (primal) dispatch solution as well as prices (dual solution). An important feature that such a framework should incorporate is decentralized decision-making. This motivates the introduction of a residual supply function, which we include to the market clearing platform in order to parallelize the computational effort and preserve the privacy of the different entities. We assess the quality of

the implemented platform on national-scale instances of Italy and Denmark. We show that our decentralized framework manages to return primal and dual solutions in less than five minutes for the Danish test case (+3,000 bus test system) while a centralized approach fails to meet this target.

Chapter 5 presents a game theoretical model of two decentralized coordination schemes. The different interpretations of the schemes motivate us to consider two simultaneous non-cooperative games and one Stackelberg game. By deriving the equilibria for the three games on an illustrative example, we demonstrate the potential inefficiencies of each of the coordination schemes.

Acknowledgments

As solitary as the work of a doctoral student is, I could not have graduated from a PhD – finally! – wihtout the people who surrounded me during these last five years.

This journey indeed started five years ago with my first meeting with Professor Anthony Papavasiliou. I came to Louvain-la-Neuve from Paris for a one-hour meeting. I don't really remember this meeting but I knew that I wanted to work with him. After a period of adaptation which mostly served at finding out the best way to collaborate, his support and expertise were of paramount help in my progress. Most importantly, his availability – considering his busy schedule – makes the difference. His door (and more recently his Skype account) is always open to discussions which significantly shorten the unavoidable periods of doubt in research. The recognition that he recently received is the best way to understand how lucky I have been to work with him. I could not have done it without you Anthony, thank you very much.

I wish to thank the members of my jury, Prof. Mathieu Van Vyve, Prof. François Glineur, Prof. Burak Kocuk and Dr. Hélène Le Cadre, for their careful review of my doctoral work. The overall quality of the manuscript benefited from your feedback.

Thanks to Anthony, I also had the chance to be a part of fruitful collaborations. I would like to thank Dr. Hélène Le Cadre for her expertise and patience in our joint work. I would also like to extend my gratitude to Pr. Ion Necoara and Pr. Quoc Tran-Dinh for their rigor and advice. I also had the chance to spend four months at the Theory Division of Los Alamos National Lab in 2019. Many thanks to Dr. Sidhant Misra and Dr. Deepjyoti Deka for making it possible and for welcoming me as part of the team. I would also like to thank everyone who made the experience personally complete – Jonas, Bailey, Kate and Nick to name a few. I had the opportunity to work in the scope of H2020 projects, namely SmartNet and FEVER. I would also like thank the people at N-SIDE and Henex for their valuable input as close partners in these projects.

I also want to acknowledge people who punctually intervene, advise or comment on my work. The proximity to Pr. Yves Smeers' office has been very convenient as soon as a question came up during our meetings with Anthony and Yves' input and thoughts are always valuable. A special thanks also to Dr. Andreas Erhenmann, Dr. Guillaume Erbs and Dr. Ibrahim Abada for introducing me to Anthony and for following some of this work, I am looking forward to collaborating with you in the future. I also had the chance to help supervising three talented former students, Taku, Marius and Alexandre, for their master theses which were related to some of the topics covered in the manuscript.

The atmosphere at CORE had also a very positive impact on my life as a PhD student. Lunch breaks have always been an opportunity to talk about

research or completely unrelated subjects. We almost started at the same time with Gilles and I'm really happy to see us both graduate at the same time as well. I also want to thank the other students of the energy group – 'energy group' is the nice brand name to refer to Anthony's students I guess – and namely Ignacio, Yuting, Céline, Loïc, Quentin, Daniel, Jacques, Jehum and Nicolas – in the order of arrival in the group, even though I am not sure that I have it right. I had great moments with you and it was good sharing this experience with you. We also had another member during a few months with Simone adding a pleasant Italian touch to our already cosmopolitan group. Many thanks also to my two office mates Cyrille and Thomas, I really enjoyed sharing my office with you. Everything not directly related to research at CORE would not have been easy to handle without the precious help of Catherine so I also wanted to thank you for your patience and your enthusiasm each time I came asking for information.

As demanding as doctoral research can be, it is impossible to find the motivation and focus if you are not able to take your mind off research. Thanks to my friends and family, I could make the most of my spare time. I won't be able to properly thank them in English, so let me write this part in French.

Commencer mon doctorat à Louvain-la-Neuve m'a donné la chance de revenir vivre à Bruxelles. J'en étais déjà très heureux mais beaucoup de personnes ont rendu ce 'retour' génial. Les premiers mois ont été rythmés par des rencontres qui sont devenues des amitiés qui perdureront encore longtemps. Merci à Aurore, Flavien, Jack, Maxime, Thibault, Victor et plus généralement 'La Banda' pour ces moments inoubliables à discuter, rigoler, faire la fête, ces moments de vie tout simplement. J'en ai oublié deux qui méritent toute ma reconnaissance et ma gratitude. Pour votre bonne humeur, pour les soirées/weekends/petits voyages, pour les nombreuses discussions qu'on a eues, pour votre soutien, pour les matches du FC Metz et les quelques virées milanaises, merci pour tout Valentin et Gaïa. Les matches de foot du samedi après-midi – et les 3èmes mi-temps – ont aussi contribué à mon épanouissement et pour cela merci le Bru Bru Bru. Pour ne pas faire de jaloux, je ne nommerais qu'un seul joueur de l'équipe et mon plus ancien ami, merci Arthur – *hermano*.

J'ai aussi la chance d'avoir pu compter sur mes amis en dehors de Bruxelles. Alexis, Jérémy, Romain, Romuald, cela fait plus de 10 ans maintenant que 'les artistes' existent et vous savez qu'ils existeront encore longtemps, je vous considère comme des membres de ma famille, merci pour tout. Dans la même lignée, merci également pour les bons moments et les belles vacances à Clémentine, Isabelle, Juliette, Mathilde et Zélia. Il est toujours bon de pouvoir parler librement du doctorat avec des amis qui traversent la même aventure, Kimia et Mahmoud merci pour cela et pour tellement plus, merci pour votre amitié sans faille, j'espère vous revoir très vite – bon courage pour la fin de thèse en attendant. Un autre ami que je n'oublie pas, Jacques, merci pour ton soutien et vivement le prochain city-trip dès qu'on en aura l'occasion.

Je ne serais sûrement pas en train d'écrire ces lignes sans l'amour et la bienveillance des familles Mezghani et Elloumi qui m'ont vu grandir. Je pense à vous, mes grands-parents, mes oncles, mes tantes, les cousins et les cousines – petite dédicace à Kaïs, c'était top de t'avoir avec nous pendant deux ans. Évidemment, merci Papa et Maman pour tout votre soutien et votre amour, je n'ai jamais manqué de rien grâce à vous. Merci pour l'équilibre parfait entre football et mathématiques – je vous laisse deviner le domaine de prédilection de chacun. Mes études sont enfin finies et j'espère vous rendre fiers. Merci aussi à ma petite soeur Lina, je te taquine souvent en disant que tu suis mes traces mais tu me surpasses depuis bien longtemps maintenant, et le meilleur reste à venir.

I am lucky to be surrounded and supported professionally and personally by everyone I've mentioned. You definitely made me go through the PhD adventure, and you played a role in the degree I will be receiving. My favorite club's motto is: *You'll never walk alone*; and I am thankful for everyone who walks by my side.

The research leading to this dissertation has received funding from the European Union's Horizon 2020 research and innovation program under grant agreements No 691405 (SmartNet) and No 864537 (FEVER). The author also acknowledges the financial support of ENGIE-Electrabel. This dissertation reflects only the author's view and the different funding entities are not responsible for any use that may be made of the information it contains.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The European Commission recently updated the terms of the European Green Deal [Com19] for the European Union (EU). Aiming at achieving carbon neutrality by 2050, the EU decided to increase the 2030 goal to 55% greenhouse gas emissions reductions from 1990 levels in December 2020 [Com20b] (the original target for 2030 was 40% and was set in 2014). This intensification of the objectives follows the willingness of the Commission to rapidly reduce the greenhouse gas emissions of all the members of the EU, initiated by the Paris Agreements and the original Green Deal.

One important source of carbon emissions is the production of energy, and in particular electricity. The EU heavily targets shifting electricity production to minimal greenhouse gas emission technologies, namely renewable energy sources (RES). To this end, the EU sets a goal of achieving a share of 38-40% from renewable resources in the European electricity mix by 2030. In 2018, the share of renewable resources represented 18% of final energy consumption, which indicates the considerable effort that is required until 2030 [Com20b]. Nevertheless, the energy sector, and particularly the electricity sector, is rapidly changing: between 2010 and 2018, solar and wind capacity grew from 110 GW to 261 GW, highly accelerated by European policy incentives and a cost decline of wind and solar electricity by 75% [Com20b].

RES technologies include solar photovoltaics (PV), wind power, hydroelectric energy and biomass. Electric vehicles and batteries are usually complementary to RES because they support the broader use of RES to prevent greenhouse gas emissions. A significant share of RES is connected to distribution networks, which implies changes in the paradigm for operating power systems. Electricity networks and markets have commonly been organized in the past with a clear distinction among two portions of the network:

(i) **Transmission network**. This is the part of the network where large generation units (responsible for most of the greenhouse gas emissions) as well as industrial consumers are typically connected. The Transmission System Operator (TSO) is responsible for operating the transmission system securely. In the absence of RES, the production of electricity typically only originates from the transmission system.

(ii) **Distribution network**. The Distribution System Operator (DSO) ensures the delivery of electricity to residential and commercial loads. In the past, the distribution network has typically consumed power, and the DSO has had a less proactive role in the optimal coordination of resources in the system.

The proliferation of RES in the distribution network is shifting the paradigm towards the coordination of both networks in the production and delivery of electricity. Numerous aspects of the implications of this transformation need to be accounted for carefully, including: (i) uncertainty and intermittency of RES; (ii) complexity and detailed modeling of the distribution networks; (iii) to what extent the role of the DSO should be broadened; (iv) coordination of TSO and DSO operations; (v) determining electricity prices.

This dissertation explores different aspects of the evolution of electricity markets, with an emphasis on organizing networks and markets that include both transmission and distribution systems in the scope of European electricity market design. Chapter 2 is divided into two parts. The first part focuses on the consideration of methods for solving non-convex optimization problems, which are applied to one of the most common problems in power systems, the alternating current optimal power flow (AC-OPF). The focus on AC-OPF is motivated by the physical characteristics of distribution networks, which requires more detailed modeling than the linear approximations of power flow that are employed in transmission network modeling. The other aspect treated in this chapter concerns the uncertainty of RES. The second part of chapter 2 proposes a practical framework for solving a stochastic version of AC-OPF. Chapter 3 discusses different coordination schemes for integrating the operations of DSOs and TSOs. Chapter 4 implements a real-time market clearing platform that integrates both transmission and distribution networks. This contribution aims at demonstrating a functional solution for clearing the market in large-scale networks. The market clearing involves computing dispatch decisions and consistent prices. Finally, chapter 5 aims at assessing several TSO-DSO coordination schemes using game theory.

The present chapter provides the necessary background on transmission-distribution (T&D) coordination in a European context (section 1.2), optimization (section 1.3) and optimal power flow (section 1.4) before presenting the structure of the dissertation (section 1.5).

1.2 T&D Coordination in European Electricity Markets

1.2.1 The Need for T&D Coordination

The deployment of renewable resources and flexible resources in medium and low-voltage distribution systems has generated an interest by the academic community and practitioners to design and implement "flexibility" platforms [CNH$^+$16] in recent years. The potential benefits of such platforms are numerous. They can support an increased deployment of distributed renewable supply (e.g. rooftop solar), safeguard the distribution network and postpone distribution network expansion, and mobilize demand-side flexibility, which in itself produces numerous short-term operational efficiencies and long-term benefits in terms of generating robust investment signals for the market [JT07].

On a high level, such flexibility platforms aim at transposing the advanced optimization functions that coordinate resources at the high-voltage grid down to the medium (and eventually low) voltage network. The interest is in accommodating distribution network constraints (voltage limits, reactive power flows, power losses, thermal limits of lines, and so on) in the optimization of system operation. This challenge necessarily increases the scale of the resulting coordination problem since resources need to be accounted for at the scale of individual distribution nodes.

The need for further TSO-DSO coordination is also stressed by the European Commission and highlighted in regulations such as the Clean Energy Package or the Network Codes. The former stresses the new important role that the DSO has to play in the upcoming energy landscape and emphasizes its "*active role*" in grid operation. For instance, Article 53.1 of the European electricity regulation states: "*Distribution system operators shall cooperate with transmission system operators in planning and operating their networks. In particular, transmission and distribution system operators shall exchange all necessary information and data regarding the performance of generation assets and demand side response, the daily operation of their networks and the long-term planning of network investments, with the view to ensure the cost-efficient, secure and reliable development and operation of their networks*" [Com17]. The fact that market mechanisms are essential to ensure this coordination is also highlighted by the regulation. A more extensive discussion of regulatory incentives behind TSO-DSO coordination in the EU can be found in [LÁT19, Section 2].

1.2.2 Ongoing Initiatives

The research and academic community have responded forcefully to the challenge of increasing the coordination of transmission and distribution system

operators. Within Europe, a number of related projects are ongoing or have recently been completed:

(i) Enera [SM20] is a project funded by the German ministry of Economic Affairs and Energy. The main objective of Enera is to operate an exchange-based flexibility market for grid congestion management, thereby reducing the need for curtailment of renewable generation.

(ii) GOPACS [SM20] is owned and operated by the Dutch TSO (TenneT) and four Dutch DSOs. GOPACS acts as an intermediary between the needs of network operators and market platforms, concretely the Dutch flexibility / market platform named ETPA.

(iii) NODES [SM20] has been launched in Norway (Norflex) and Germany. Whereas in Norway the problem that is being addressed is downward flow to loads, in Germany the desire is to reduce curtailment of wind. NODES is currently operated on a continuous basis as an intraday market.

(iv) Piclo Flex [SM20] acts as a bulletin board, where DSOs post their customized and localized needs for flexibility (essentially a volume, a location, an up / down direction and a period, as well as certain technical characteristics). Asset owners then respond to such "tenders".

(v) The Cornwall local energy market [EGG17] is a local market project developed in the Cornwall region in the UK. The Cornwall local energy market allows both the TSO and the DSO to trade both reserve and energy under a closed-gate auction mechanism which is conducted both on a day-ahead and intra-day basis and that includes detailed network constraints of the DSO.

(vi) Soteria is a local market that is being put in place by Fluvius (a Belgian DSO) and Elia (the Belgian TSO). The market platform developed in the Soteria project (`ioenergy.eu/soteria`) is aiming at unlocking more residential flexibility from the DSO grid to be used by the TSO for balancing, while respecting DSO network constraints.

(vii) CoordiNet [MGASM$^+$20] is an EU-funded project centered around TSO-DSO cooperation which targets the development of multiple pilot markets in Spain, Greece and Sweden. As an example, in the local market being developed in Spain for the DSO from the perspective of local congestion management, the DSO, who is the sole buyer in this platform, bids its network constraints and flexibility needs in the market, which is organized as a closed-gate auction in the day-ahead and intraday time frame.

(viii) SmartNet [MRS$^+$17], another EU project, is the predecessor of Coordinet. This project aims at integrated balancing and congestion management in

real time, and investigates various TSO-DSO coordination mechanisms for achieving this goal.

(ix) EU-SysFlex (`eu-sysflex.com`) is an H2020 project which aims at providing a long-term plan for integrating new technologies capable of furnishing flexibility in the large-scale pan-European market.

(x) DOMINOES (`dominoesproject.eu`) is a European project which demonstrates how DSOs can actively manage the network in order to handle power balance in emerging future configurations such as microgrids, ultra-distributed generation and energy independent communities.

(xi) Ecogrid [EGG17] is a Danish project which took place on Bornholm island. The goal of the project was to test a regional balancing market where DERs and flexible consumers are encouraged to participate by adapting to fluctuations of electricity prices based on local congestion.

These projects cast a wide net of operational objectives. They are all centered around relatively short-term operational time scales (typically day-ahead, intraday, or real time), and involve some extent of transmission and distribution system coordination. The variety of market designs that is proposed in these projects is also relatively wide. In Table 1.1, we summarize the scale on which these projects are conducted as well as their main focus.

1.2.3 European Electricity Markets

The diversity of projects that are referenced in the previous section indicates a wide range of approaches to the same fundamental problem, which is one of increasing the coordination of flexible resources in both the transmission as well as distribution system. A fundamental question that is either explicitly or implicitly motivating these flexibility projects is how the proposed market design can be integrated to existing EU market design [Com20a]. In this section we briefly overview the structure of the EU market, in order to better frame the problem for the remainder of the dissertation.

The EU short-term market, like numerous international markets, is organized as a sequence of day-ahead and intraday forward markets followed by a real-time balancing market. This sequence is represented in Figure 1.1:

- The **reserve** or **balancing capacity** market is operated before, simultaneously, or after the day-ahead energy market (depending on the specific reserve product and the country in which it is traded). This market is operated by the entities tasked with maintaining the secure operation of the network, namely TSOs. The goal of the balancing capacity market is to guarantee that sufficient overhead capacity is scheduled for the following day in order to ensure that the system can operate securely even

Table 1.1: Ongoing European projects dealing with flexibility markets, RES integration and T&D coordination.

Project	Scale	Main focus?		
		Flexibility market	RES integration	T&D coordination
Enera	National: Germany	✓	✓	
GOPACS	National: Netherlands	✓		✓
NODES	National: Norway, Germany		✓	
Piclo Flex	National: UK		✓	
Cornwall	Regional: Cornwall (UK)			✓
Soteria	National: Belgium		✓	✓
CoordiNet	European: Spain, Greece, Sweden	✓		✓
SmartNet	European: Italy, Spain, Denmark			✓
Eu-SysFlex	European		✓	
DOMINOES	European		✓	✓
Ecogrid	Regional: Bornholm (Denmark)	✓	✓	

if imbalances or outages of system components (generators or transmission lines) occur in real time. Various balancing capacity products are traded in the balancing capacity market, including (in order of their so-called Full Activation Time, i.e. the speed by which they can respond to system imbalances) frequency containment reserve, automatic frequency restoration reserve, manual frequency restoration reserve, and replacement reserve.

- The **day-ahead** market is a pan-European auction based on uniform pricing principles which is operated by power exchanges, also referred to as Nominated Electricity Market Operators (NEMOs). The day-ahead market allows participants to place buy or sell bids up to a certain time (around 12pm) the day before the dispatch. The market clears at once for the entire 24 hours of the following day. The market is cleared in the afternoon of the day before operations by the NEMOs in a common pan-European market platform. The resulting market clearing determines energy prices and accordingly accepts or rejects the inserted bids by attempting to match bids so as to maximize economic welfare.

- The **intraday** market permits participants to adjust and refine their position as real time approaches and more accurate system conditions are revealed. The intraday market is also operated by NEMOs, and functions as a continuous trading platform where buy and sell bids are matched as they arrive to the platform. Bids are submitted for a resolution of 15 minutes. Trading in the intraday market concludes 30 minutes before delivery time.

- The **real-time** market, also called **balancing** market is the last market for trading energy before delivery. This market adjusts the net positions of the participants by activating the reserves. It is handled by the TSO who is in charge of securely operating the network. The balancing energy market is connected to the balancing capacity market, in the sense that those reserves which have been accepted successfully in the reserve market for delivering a certain amount of capacity to the system operator are required to bid at least that amount of energy in the balancing energy market. It is in this sense that these resources offer reserve to the system: if cleared for balancing capacity in the reserve market, they commit to be available to provide balancing energy in real time. This is the market mostly considered in this dissertation, in Chapters 3, 4 and 5. Despite regional specificities, the EU balancing market is moving towards full integration following EU legislation and a number of EU-wide integration projects. In the foreseen integrated market, the balancing platforms will be operated as uniform price auctions for energy, and it is with these provisions in mind that we conduct our analysis in the present work.

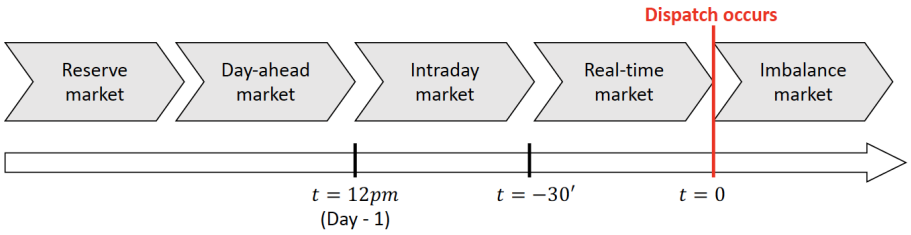


Figure 1.1: Chronology of the European electricity markets.

Apart from describing the chronology of market operations, it is useful to clarify the products that are traded in each market. Before proceeding, we introduce some further terminology related to the EU market. Balancing Service Providers (BSPs) in this thesis essentially refer to owners of reserve assets, i.e. entities that can submit price-elastic bids in real time for buying or selling energy in real time. Balancing Responsible Parties (BRPs) in this thesis refer to price-inelastic buyers or sellers of real-time energy, i.e. entities that

experience a real-time energy imbalance that the system operator is required to balance out by mobilizing reserves in the system.

Having described the full range of products in the market, we now clarify what products each market trades. The balancing energy market trades real-time energy between BSPs and BRPs, as well as transmission capacity from the system operator to market participants. This market is the focus of the present work. Note that, as the balancing energy market clears every fifteen minutes, the amount of time afforded for actually performing the matching of bids and pricing is quite limited. This concern is explicitly accounted for in the subsequent material. The intraday and day-ahead energy market execute the exact same function as the balancing energy market, but in a forward setting, i.e. so as to allow market agents to mitigate risk by trading energy in advance of real time. The balancing capacity market trades reserve / balancing capacity between BSPs and the TSO.

1.3 Preliminaries on Optimization

Optimization is a useful tool to model and solve problems appearing in power networks and electricity markets. A continuous optimization problem consists of minimizing a certain objective subject to a certain number of constraints and is written as follows:

$$u^* = \min_{x \in \mathbb{R}^d} \; C(x) \tag{1.1a}$$

$$s.t. \quad f_i(x) \leq b_i, \; \forall i = 1, \ldots, N. \tag{1.1b}$$

where x represents the *decision variables* of the problem, C is the *objective function* (or *cost function*), u^* is the *optimal value*, $(f_i)_{i=1,\ldots,N}$ are the *constraint functions* and $(b_i)_{i=1,\ldots,N}$ are the *constraint bounds*. In this section, we assume that the objective is linear, i.e. $C(x) = \sum_{j=1}^{d} (C_{j,1} x_j + C_{j,0})$, and that $(f_i)_{i=1,\ldots,N}$ are continuous. Only certain classes of the general problem (1.1) can be solved. If functions $(f_i)_{i=1,\ldots,N}$ are convex, problem (1.1) is a convex optimization problem. Convex optimization problems have global optimality guarantees and most of them are efficiently solved with powerful solvers. Non-convex optimization problems are much harder to solve. We explore the relevant classes of optimization problems for this dissertation in the next subsections.

1.3.1 Linear Programming

A linear program (LP) is the elementary optimization problem. Indeed, LPs consider the constraint functions as linear:

$$f_i(x) = \alpha_i^\top x + \beta_i,\ \alpha_i, \beta_i \in \mathbb{R}^d,\ \forall i = 1, \ldots, N$$

where superscript $\top$ represents the scalar product on $\mathbb{R}^d$. LPs are efficiently solved nowadays and most of the solution methods rely on the simplex algorithm [DOW+55] or interior point methods [LMS94].

1.3.2 Second Order Cone Programming

Another class of optimization problem which is efficiently solved by current solvers is second order cone programs (SOCP). A second order cone constraint function has the following form:

$$f_i(x) = ||A_i x + e_i||_2 + \alpha_i^\top x + \beta_i,\ A_i \in \mathbb{R}^{a_i \times d}, e_i \in \mathbb{R}^d,\ \forall i = 1, \ldots, N$$

where $||.||_2$ is the Euclidean norm. The format of second-order cone constraints is particularly useful because it includes constraints such as:

$$y^2 + z^2 \leq S \iff \left\| \begin{pmatrix} y \\ z \end{pmatrix} \right\|_2 \leq \sqrt{S}$$

$$\begin{cases} y^2 + z^2 \leq w_1 w_2, \\ w_1, w_2 \geq 0 \end{cases} \iff \begin{cases} \left\| \begin{pmatrix} y \\ z \\ (w_1 - w_2)/2 \end{pmatrix} \right\|_2 \leq (w_1 + w_2)/2 \\ w_1, w_2 \geq 0 \end{cases}$$

where y, z, w_1, w_2 are variables and S is a parameter. These two types of constraints are particularly useful to model the physics of a power network. Most commercial solvers (like CPLEX, Mosek or Gurobi) implement a variation of an interior point method, namely *homogeneous and self-dual algorithm* [ART03], to solve SOCPs.

1.3.3 Non-Convex Optimization

Non-convex optimization problems consider the case where there is no assumption on $(f_i)_{i=1,\ldots,N}$. Nevertheless, an efficient interior-point method, the *interior point line search filter method* [WB06], have been implemented when considering the case where $(f_i)_{i=1,\ldots,N}$ are twice continuously differentiable. Interior point methods guarantee convergence to a local minimum or an infeasible point, which is unsatisfactory in general when solving (feasible) optimization problems. Also, even if solvers are becoming more and more robust and

fast [WB06], these methods require more computation time and face more numerical instability than LPs or SOCPs. These issues can partially be resolved if the solver is carefully initialized by an accurate approximate solution. To solve this type of problems, we usually first focus on approximations or relaxations of the problem. The idea of approximations and relaxations also holds for convex problems.

1.3.3.1 Approximations

Optimization problems allow for modeling practical problems and their applications come from transport, telecommunication, electricity markets, finance etc. If problem (1.1) is too hard to tackle, practitioners might be in favor of solving:

$$u^*_{\text{approx}} = \min_{x \in \mathbb{R}^d} \; C(x) \tag{1.2a}$$

$$s.t. \quad \tilde{f}_i(x) \leq b_i, \; \forall i = 1, \ldots, N. \tag{1.2b}$$

where $(\tilde{f}_i)_{i=1,\ldots,N}$ are approximations of $(f_i)_{i=1,\ldots,N}$ that have an easier form to handle (typically linear or second-order cone constraints). The quality of the approximation will only depend on its accuracy compared to the original problem. Nevertheless, in general, no link can be established between the solution of (1.1) and (1.2).

1.3.3.2 Relaxations

Another way of approaching a difficult optimization problem is through relaxations. There are several ways of relaxing an optimization problem, but the main idea is to consider a larger feasible space. The feasible space of (1.1) is:

$$\mathcal{F} = \{x \in \mathbb{R}^d \mid \forall i = 1, \ldots, N, f_i(x) \leq b_i\}$$

By considering $\mathcal{F} \subseteq \mathcal{F}_{\text{relax}}$, a relaxation of (1.1) is written as:

$$u^*_{\text{relax}} = \min_{x \in \mathbb{R}^d} \; C(x) \tag{1.3a}$$

$$s.t. \quad x \in \mathcal{F}_{\text{relax}} \tag{1.3b}$$

For example, if functions $(f_i)_{i=1,\ldots,N}$ are non-convex, one way of relaxing (1.1) is replacing (1.1b) by:

$$\underline{f}_i(x) \leq b_i, \; \forall i = 1, \ldots, N.$$

where $\underline{f}_i(x) \leq f_i(x), \; \forall x \in \mathbb{R}^d, \; \forall i = 1, \ldots, N$. $(\underline{f}_i)_{i=1,\ldots,N}$ would typically be chosen as linear, or at least convex.

Solving a relaxation has the nice property of providing a lower bound of the optimal value:

$$u^*_{\text{relax}} \leq u^*$$

In the next section, we will see how optimization problems allow for detailed modeling of power networks. Since the classical alternating current optimal power flow (AC-OPF) problem is non-convex, several approaches based on approximations and relaxations enable attacking this difficult problem.

1.4 Preliminaries on Optimal Power Flow

In this section we discuss formulations of the Alternating Current Optimal Power Flow (AC-OPF) that are employed in the dissertation, as well as a direct current (DC) approximation, a trigonometric reformulation, a quadratic relaxation and a second order cone programming (SOCP) relaxation.

1.4.1 The Power Flow Problem

The Power Flow (PF) equations emerge in power systems and allow representing the physical properties, such as Kirchhoff's Law and Ohm's Law, on a power network. In a general setting, referred to as alternating current (AC), the basic variables of the PF equations are (P, Q, v, θ): real power injection, reactive power injection, voltage magnitude and voltage angle. Even if they are not necessary for basic formulations of PF equations, we also introduce real (resp. reactive) power flow f^p (resp. f^q) for modeling purposes. The set of decision variables is then $x = (P, Q, v, \theta, f^p, f^q)$. We denote by $\mathcal{N} = (\mathcal{B}, \mathcal{L})$ a power network, where $\mathcal{B}$ is the set of buses and $\mathcal{L}$ is the set of directed lines ($\mathcal{L}^R$ being the set of reverse directed lines [CHVH15]). The parameter Y, defined as $Y_{ik} = G_{ik} + jB_{ik},\ \forall (i,k) \in \mathcal{L} \cup \mathcal{L}^R$ ($j^2 = -1$), is the admittance matrix, which relies on the conductance G and the susceptance B. The admittance matrix reflects the properties of the network. The AC PF equations are then written

as follows:

$$P_i = \sum_{j \in \delta(i)} f^p_{ij} + G_i v_i^2, \qquad \forall i \in \mathcal{B} \tag{1.4a}$$

$$Q_i = \sum_{j \in \delta(i)} f^q_{ij} - B_i v_i^2, \qquad \forall i \in \mathcal{B} \tag{1.4b}$$

$$f^p_{ij} = -G_{ij} v_i^2 + v_i v_j (G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j)), \quad \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R \tag{1.4c}$$

$$f^q_{ij} = B_{ij} v_i^2 + v_i v_j (G_{ij} \sin(\theta_i - \theta_j) - B_{ij} \cos(\theta_i - \theta_j)), \quad \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R \tag{1.4d}$$

Here, $\delta(i) = \{j \in \mathcal{B} \mid (i,j) \in \mathcal{L} \cup \mathcal{L}^R\}$ defines the set of buses directly connected to bus $i \in \mathcal{B}$. Equations (1.4) is a system of non-linear equations of $4|\mathcal{B}|$ variables and $2|\mathcal{B}|$ equalities (by substituting (1.4c)–(1.4d) in (1.4a)–(1.4b)). We usually solve the power flow equations by specifying the values of $2|\mathcal{B}|$ *control* variables. This is the power flow problem. The choice of the control variables is based on the representation of an ordinary power network. We distinguish different types of buses:

- PV buses (set $\mathcal{PV}$): also called *generator* bus, this type of bus specifies the real power generation P and the voltage magnitude v.
- PQ buses (set $\mathcal{PQ}$): commonly called *load* bus, the PQ bus enforces real and reactive power consumption P and Q.
- Slack bus (denoted *Slack*): this single bus is able to generate or consume power. It typically imposes voltage magnitude to $v = 1$ and voltage angle to $\theta = 0$.

(P, v) for PV buses, (P, Q) for PQ buses and (v, θ) for the slack bus represent the control variables when solving the PF problem and the rest of the variables are seen as the *state* variables. This partition of the buses changes the system (1.4) to a $2|\mathcal{B}|$-variables and $2|\mathcal{B}|$-equalities system of non-linear equations. The PF problem has largely been studied [SA74, Sto74, SM79] and is solved using the Newton-Raphson method for example [Sto74].

Another application of the power flow equations (1.4) does not differentiate state and control variables. If we do not fix some of the free variables, then we can optimize them in order to minimize a certain objective, which is typically operating cost, but also respecting certain bounds on the control and state variables (for example, generation capacity or flow limits). This gives rise to the Optimal Power Flow (OPF) problem.

1.4.2 Optimal Power Flow Formulations

1.4.2.1 Classic OPF Formulation

The classical AC-OPF problem, formulated originally by [Car62], aims at dispatching generating units on a network in order to satisfy demand, while respecting physical properties (Kirchhoff's Law and Ohm's Law) and thermal and operational constraints. We assume that generators are dispatched with an objective of minimizing active power generation costs. In this dissertation, at a certain bus $i \in \mathcal{B}$, we assume that the active power injection is the difference of the active power produced by generators connected to bus i (this set is denoted $\mathcal{G}_i$) and the active power demand D_i^p, assumed to be known and fixed: $P_i = \sum_{g \in \mathcal{G}_i} p_g - D_i^p$. The same assumption is made for reactive power. The resulting classical AC-OPF problem is formulated as follows:

$$\min_{x} \quad \sum_{g \in \mathcal{G}} C_g(p_g) \tag{1.5a}$$

$$s.t. \quad \sum_{g \in \mathcal{G}_i} p_g - D_i^p = \sum_{j \in \delta(i)} f_{ij}^p + G_i v_i^2, \ \forall i \in \mathcal{B} \tag{1.5b}$$

$$\sum_{g \in \mathcal{G}_i} q_g - D_i^q = \sum_{j \in \delta(i)} f_{ij}^q - B_i v_i^2, \ \forall i \in \mathcal{B} \tag{1.5c}$$

$$f_{ij}^p = -G_{ij} v_i^2 + v_i v_j (G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j)), \ \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R \tag{1.5d}$$

$$f_{ij}^q = B_{ij} v_i^2 + v_i v_j (G_{ij} \sin(\theta_i - \theta_j) - B_{ij} \cos(\theta_i - \theta_j)), \ \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R \tag{1.5e}$$

$$\underline{v}_i \leq v_i \leq \overline{v}_i, \ \forall i \in \mathcal{B} \tag{1.5f}$$

$$(f_{ij}^p)^2 + (f_{ij}^q)^2 \leq S_{ij}^2, \ \underline{f}_{ij}^p \leq f_{ij}^p \leq \overline{f}_{ij}^p, \ \underline{f}_{ij}^q \leq f_{ij}^q \leq \overline{f}_{ij}^q, \ \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R \tag{1.5g}$$

$$p_g^2 + q_g^2 \leq \overline{pq}_g^2, \ \underline{p}_g \leq p_g \leq \overline{p}_g, \ \underline{q}_g \leq q_g \leq \overline{q}_g, \ \forall g \in \mathcal{G} \tag{1.5h}$$

Here, $C_g(p_g) = C_{g,1} p_g + C_{g,0}, \ (C_{g,1}, C_{g,0}) \in \mathbb{R}^2, \ \forall g \in \mathcal{G}$ is linear. Constraints (1.5b)–(1.5e) refer to the power flow equations. Constraints (1.5f)–(1.5h) are the technical constraints that might be violated when only solving PF. Constraint (1.5f) describes the voltage limits at each node. Constraints (1.5g) refer to (i) apparent power line limits, (ii) real power limits, and (iii) reactive power limits. In the same spirit, for each generator, constraints (1.5h) impose apparent injection limits as well as lower and upper limits on real and reactive injections. Even if all the constraints of (1.5g) and (1.5h) are not always considered when introducing AC-OPF [LL11, FL13, GLTL14, MH+19], we

make formulation (1.5) as general as possible which is adapted to the SmartNet instances that will be used in chapter 4.

In order to have a more concise formulation of the problem, we introduce further notations. Constraints (1.5d)–(1.5g) relate to operational constraints that are handled by the network operator and constraints (1.5h) refer to generation capacity constraints. For the sake of conciseness, we therefore introduce the set of operational constraints $\mathcal{OC}$, and the set of generation constraints $\mathcal{GC}_g$ which are defined as:

$$\begin{aligned} \mathcal{OC} &= \{(v, \theta, f^p, f^q) \text{ satisfying } (1.5\text{d}) - (1.5\text{g})\} \\ \mathcal{GC}_g &= \{(p_g, q_g) \text{ satisfying } (1.5\text{h})\} \qquad\qquad \forall g \in \mathcal{G} \end{aligned}$$

We will also use a compact notation for the power flow constraints (1.5b)–(1.5c) by introducing functions $F_i,\ i \in \mathcal{B}$ as follows:

$$F_i(x) = 0,\ \forall i \in \mathcal{B} \Longleftrightarrow x = (p, q, v, \theta, f^p, f^q) \text{ satisfies } (1.5\text{b}) - (1.5\text{c}).$$

The concise formulation of AC-OPF is then:

$$\begin{aligned} \text{OPF}_{\text{AC}} : \min_x \quad & \sum_{g \in \mathcal{G}} C_g(p_g) && && (1.6\text{a}) \\ s.t. \quad & F_i(x) = 0, && \forall i \in \mathcal{B} && (1.6\text{b}) \\ & (v, \theta, f^p, f^q) \in \mathcal{OC} && && (1.6\text{c}) \\ & (p_g, q_g) \in \mathcal{GC}_g, && \forall g \in \mathcal{G} && (1.6\text{d}) \end{aligned}$$

Problem (1.6) (and equivalently (1.5)) is a non-convex optimization problem (due to constraints (1.5b)–(1.5e)) and relaxations and exact approaches have largely been studied in the literature [MAEH99, MEHA99, FSR12, MH+19]. In the thesis, we first introduce the classical DC approximation (subsection 1.4.2.2). Based on [MH+19], relaxations of AC-OPF can be split into three categories: linear relaxations, second-order cone relaxations and semi-definite relaxations. If one were to summarize the relationships between these three categories, one would say that linear relaxations are computationally easy to solve but provide less accurate solutions whereas semi-definite relaxations favor accuracy over computational burden. Since second-order cone relaxations lie in between these two categories and are particularly appealing on radial networks, we only consider a second-order cone relaxation of AC-OPF in the thesis. Based on a reformulation of AC-OPF (subsection 1.4.2.3), we derive a non-convex quadratic relaxation of AC-OPF (subsection 1.4.2.4) which extends to a second-order cone relaxation (subsection 1.4.2.5).

1.4.2.2 Direct Current Approximation

The direct current (DC) approximation is the most commonly used approximation of the problem. It is mostly used on transmission networks where

the voltage level renders the accuracy of the approximation acceptable. The underlying assumptions of this approximation are the following:

- $G_{ij} \approx 0,\ \forall (i,j) \in \mathcal{L}$
- Angle differences are small: $\sin(\theta_i - \theta_j) \approx (\theta_i - \theta_j),\ \forall (i,j) \in \mathcal{L}$.
- $v_i \approx 1, \forall i \in \mathcal{B}$

Using (1.5d) and (1.5e), these assumptions allow for ignoring reactive power. The real power flow can then be expressed as:

$$f_{ij}^p = B_{ij}(\theta_i - \theta_j),\ \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R$$

This renders the approximation linear, as follows:

$$\begin{aligned}
\mathrm{OPF_{DC}}: \min \quad & \sum_{g \in \mathcal{G}} C_g(p_g) && && (1.7a)\\
s.t. \quad & F_i^{DC}(x) = 0, && \forall i \in \mathcal{B} && (1.7b)\\
& (\theta, f^p) \in \mathcal{OC}^{DC} && && (1.7c)\\
& p_g \in \mathcal{GC}_g^{DC}, && \forall g \in \mathcal{G} && (1.7d)
\end{aligned}$$

Here, $\mathcal{GC}^{DC} = \{p_g \text{ satisfying: } \underline{p}_g \leq p_g \leq \overline{p}_g\}$ and $\mathcal{OC}^{DC} = \{(\theta, f^p) \text{ satisfying: } f_{ij}^p = B_{ij}(\theta_i - \theta_j),\ \underline{f}_{ij}^p \leq f_{ij}^p \leq \overline{f}_{ij}^p,\ \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R\}$ are the relaxed set of generation and operational constraints in the DC approximation. We also make use of:

$$F_i^{DC}(x) = 0,\ \forall i \in \mathcal{B} \Longleftrightarrow p_i - D_i^p = \sum_{j \in \delta(i)} f_{ij}^p,\ \forall i \in \mathcal{B}.$$

1.4.2.3 The Trigonometric Reformulation

An alternative formulation of the general problem (1.5) is the trigonometric formulation introduced in [ER99, Jab06, Jab07]. The trigonometric formulation is introduced by defining the following variables:

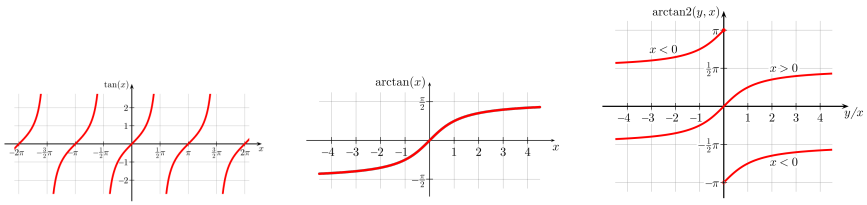
$$\begin{aligned}
& c_{ij} = v_i v_j \cos(\theta_i - \theta_j), && \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R && (1.8a)\\
& s_{ij} = -v_i v_j \sin(\theta_i - \theta_j), && \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R && (1.8b)\\
& c_{ii} = v_i^2,\ s_{ii} = 0, && \forall i \in \mathcal{B} && (1.8c)
\end{aligned}$$

By using classical trigonometric relations, the set of constraints (1.8) could be replaced by:

$$\begin{aligned}
& c_{ij}^2 + s_{ij}^2 = c_{ii} c_{jj}, && \forall (i,j) \in \mathcal{L}\\
& \tan(\theta_i - \theta_j) = \frac{v_i v_j \sin(\theta_i - \theta_j)}{v_i v_j \cos(\theta_i - \theta_j)} = \frac{-s_{ij}}{c_{ij}}, && \forall (i,j) \in \mathcal{L}
\end{aligned}$$

Since angle differences usually lie in a 2π-interval, the use of tan() is limiting in the sense that it is defined on $\mathbb{R} - \{\frac{\pi}{2} + k\pi, k \in \mathbb{Z}\}$ (Figure 1.2(a)). We would rather write $\theta_i - \theta_j$ as a function of c_{ij} and s_{ij}. Since the range of values of arctan() is $]-\pi/2, \pi/2[$ (Figure 1.2(b)), we would rather use atan2() ranging in $]-\pi, \pi]$ [KDS16] (Figure 1.2(c)) and defined as:

$$\text{atan2}(x, y) = 2 \arctan \frac{y}{\sqrt{x^2 + y^2} + x}, \forall x, y$$



(a) tan. (b) arctan. (c) arctan 2 (atan2).

Figure 1.2: Graphical comparison of tan, arctan and atan2 functions.

The set of constraints (1.8) is finally replaced by:

$$c_{ij}^2 + s_{ij}^2 = c_{ii} c_{jj}, \qquad \forall (i,j) \in \mathcal{L} \tag{1.9a}$$
$$\theta_i - \theta_j = \text{atan2}(s_{ij}, c_{ij}), \qquad \forall (i,j) \in \mathcal{L} \tag{1.9b}$$

The set of constraints (1.5b)-(1.5f) is then rewritten as:

$$\sum_{g \in \mathcal{G}_i} p_g - D_i^p = \sum_{j \in \delta(i)} f_{ij}^p + G_i c_{ii}, \qquad \forall i \in \mathcal{B} \tag{1.10a}$$
$$\sum_{g \in \mathcal{G}_i} q_g - D_i^q = \sum_{j \in \delta(i)} f_{ij}^q - B_i c_{ii}, \qquad \forall i \in \mathcal{B} \tag{1.10b}$$
$$f_{ij}^p = -G_{ij} c_{ii} + G_{ij} c_{ij} - B_{ij} s_{ij}, \qquad \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R \tag{1.10c}$$
$$f_{ij}^q = B_{ij} c_{ii} - G_{ij} s_{ij} - B_{ij} c_{ij}, \qquad \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R \tag{1.10d}$$
$$\underline{v}_i^2 \leq c_{ii} \leq \overline{v}_i^2, \qquad \forall i \in \mathcal{B} \tag{1.10e}$$

And the following trigonometric reformulation of the problem has then been proposed originally in [Jab07]:

$$\begin{aligned}
\mathrm{OPF}_{\text{AC-T}} : \min_{x} \quad & \sum_{g\in\mathcal{G}} C_g(p_g) && && (1.11a)\\
s.t. \quad & F_i^{\text{AC-T}}(x) = 0, && \forall i \in \mathcal{B} && (1.11b)\\
& (c, s, \theta, f^p, f^q) \in \mathcal{OC}^{\text{AC-T}} && && (1.11c)\\
& (1.6d) && && (1.11d)
\end{aligned}$$

The decision variables are now $x = (p, q, c, s, \theta, f^p, f^q)$. And $F^{\text{AC-T}}$ and $\mathcal{OC}^{\text{AC-T}}$ are defined as:

$$F_i^{\text{AC-T}}(x) = 0,\ \forall i \in \mathcal{B} \Longleftrightarrow (1.10a) - (1.10b)$$
$$\mathcal{OC}^{\text{AC-T}} = \{(c, s, \theta, f^p, f^q) \text{ satisfies } (1.9a) - (1.9b), (1.10c) - (1.10e), (1.5g)\}$$

Problem (1.11) is an exact formulation of AC-OPF and is then equivalent to OPF_{AC}.

1.4.2.4 The Quadratic Relaxation

Introducing $\mathrm{OPF}_{\text{AC-T}}$ allows for isolating θ variables in a single constraint (1.9b). By ignoring (1.9b) and consequently the angles θ, the quadratic relaxation of OPF is introduced as:

$$\begin{aligned}
\mathrm{OPF}_{\text{Q}} : \min_{x} \quad & \sum_{g\in\mathcal{G}} C_g(p_g) && && (1.12a)\\
s.t. \quad & F_i^{Q}(x) = 0, && \forall i \in \mathcal{B} && (1.12b)\\
& (c, s, f^p, f^q) \in \mathcal{OC}^{Q} && && (1.12c)\\
& (1.6d) && && (1.12d)
\end{aligned}$$

Here, $F^Q = F^{\text{AC-T}}$, (1.9b) is removed from $\mathcal{OC}^{\text{AC-T}}$ to define $\mathcal{OC}^Q$ and $x = (p, q, c, s, f^p, f^q)$. The relaxation (1.12) is exact when the network is radial [ZT11]. Indeed, once (1.12) is solved, recovering angles using (1.9b) implies solving a system of linear equations (with $|\mathcal{B}|$ variables and $|\mathcal{B}| - 1$ equations). Problem (1.12) is a non-convex quadratically constrained problem and serves as a basis to introduce the second order cone relaxation of OPF.

1.4.2.5 Second Order Cone Relaxation

In formulation (1.12), the only non-convex quadratic constraints are (1.9a). This equality can be relaxed to a second order cone constraint:

$$c_{ij}^2 + s_{ij}^2 \leq c_{ii}c_{jj},\ \forall (i, j) \in \mathcal{L} \tag{1.13}$$

By slightly modifying $\mathcal{OC}^Q$ to $\mathcal{OC}^{SOC}$ (one just needs to replace (1.9a) by (1.13)) and having $F^{SOC} = F^Q$, the SOCP relaxation of OPF is then formulated as follows:

$$\text{OPF}_{\text{SOC}} : \min_x \quad \sum_{g \in \mathcal{G}} C_g(p_g) \tag{1.14a}$$

$$s.t. \quad F_i^{SOC}(x) = 0, \qquad \forall i \in \mathcal{B} \tag{1.14b}$$

$$(c, s, f^p, f^q) \in \mathcal{OC}^{SOC} \tag{1.14c}$$

$$\text{(1.6d)} \tag{1.14d}$$

This formulation has been introduced by [Jab06] and is particularly useful on radial networks. Indeed, under certain assumptions [FL13, GLTL14], it is possible to guarantee that constraint (1.13) is tight and that the relaxation is exact. The idea of these assumptions is that it is possible to ensure exactness if (i) the voltage limits are not binding [GLTL14] or (ii) an unlimited amount of real and reactive power can be injected or withdrawn at each bus [FL13]. In practice, these assumptions are typically not met. Nevertheless, the SOCP relaxation on radial networks has been observed to yield feasible dispatch solutions in practice, even if the assumptions that guarantee exactness are not fulfilled [BVC20]. In the thesis, we use the SOCP solution for attempting to derive a feasible dispatch. Even if the dispatch is found to be infeasible, we use the SOCP solution for warm-starting or to derive a lower bound on the objective value.

1.4.2.6 Relations Between the OPF Formulations

The OPF formulations that we have introduced are connected as follows:

1. OPF_{AC} and $\text{OPF}_{\text{AC-T}}$ are non convex problems and are equivalent.
2. OPF_{Q} is a quadratic non-convex relaxation of $\text{OPF}_{\text{AC-T}}$.
3. OPF_{SOC} is a convex conic relaxation of OPF_{Q}.

These relationships are known and presented in [KDS16, MH+19]. Note that OPF_{DC} is a linear approximation and cannot be compared to the other formulations. These relationships are illustrated in Figure 1.3.

1.4.2.7 Computational Comparison

To illustrate the different levels of difficulty involved when trying to solve the OPF problems introduced, we display the results of 4 OPF instances solved using `PowerModels.jl` [CBS+18a] on MATPOWER instances [ZMST10]. OPF_{DC} and OPF_{SOC} are solved with Gurobi [GO19] and OPF_{AC} with IPOPT [WB06].

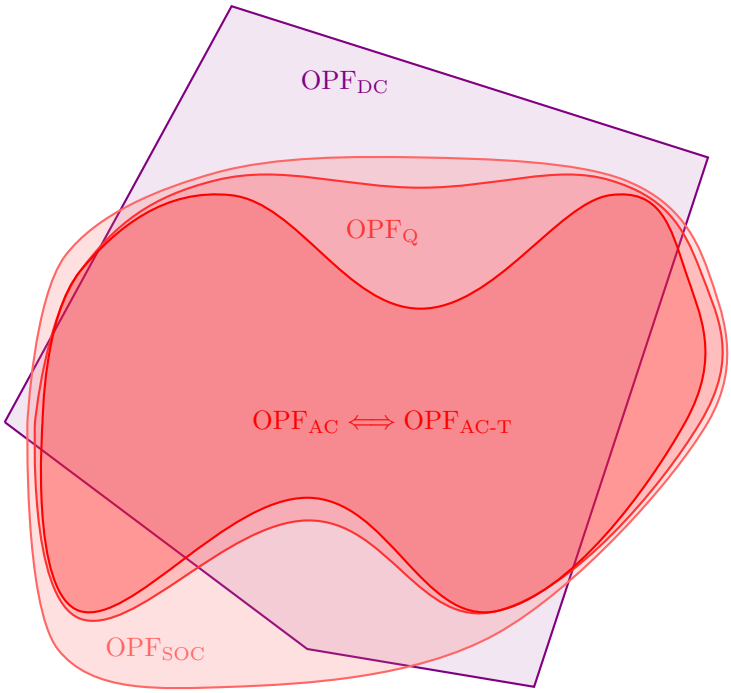


Figure 1.3: Schematic of the OPF formulations feasible spaces.

The results are shown in Table 1.2 accompanied with more information about the test cases in Table 1.3. The gap reported in Table 1.2 in columns OPF_{DC} and OPF_{SOC} shows the relative distance to the OPF_{AC} objective value. It should be highlighted that when solving OPF_{DC} and OPF_{SOC} a global minimizer is obtained while solving OPF_{AC} with IPOPT only returns a local minimum guarantee. Note also that we launch 10 runs of IPOPT with 10 different random initial vectors $(p^0, q^0, v^0, \theta^0)$ (warm-start option of the solver) and it always leads to the same objective value for the four test cases of Table 1.2. Since the run without warm-start requires less time to solve, we only report this run in Table 1.2.

We clearly see how easy it is to solve the DC approximation and to a lesser extent the SOCP relaxation. The scale of the problem also plays a role in the efficiency to solve the formulations: for `13659pegase`, OPF_{DC} solves more than 200 times faster than OPF_{AC}, and OPF_{SOC} almost 20 times faster. In terms of objective, the objective values obtained for OPF_{DC} and OPF_{SOC} are relatively close to OPF_{AC}, and we remind that only the objective value of OPF_{SOC} is a lower bound of the objective value of OPF_{AC}. The lower bounding of OPF_{SOC} and its computational performances stress the strength

Table 1.2: Comparison of the OPF formulations

	OPF_{DC}			OPF_{SOC}			OPF_{AC}	
Test Case	Objective	Gap	Time	Objective	Gap	Time	Objective	Time
`1354pegase`	$7.306e^4$	-1.4%	0.10 s	$7.401e^4$	-0.1%	2.13 s	$7.407e^4$	6.91 s
`2869pegase`	$1.324e^5$	-1.2%	0.25 s	$1.339e^5$	-0.1%	7.35 s	$1.340e^5$	29.95 s
`9241pegase`	$3.124e^5$	-1.1%	1.99 s	$3.104e^5$	-1.7%	36.20 s	$3.159e^5$	131.61 s
`13659pegase`	$3.818e^5$	-1.1%	2.63 s	$3.803e^5$	-1.5%	28.77 s	$3.861e^5$	648.24 s

Table 1.3: Number of variables and constraints of the test cases in OPF_{AC}.

Test Case	# Variables	# Constraints
`1354pegase`	11,192	16,957
`2869pegase`	25,086	37,489
`9241pegase`	85,568	123,683
`13659pegase`	117,370	146,437

of this formulation. Nevertheless, operators are interested in a feasible optimal solution of OPF_{AC} and the question of recovering feasible dispatches from approximations or relaxations is still a research question in itself and is out of the scope of this dissertation.

In practice, the use of the DC approximation is relevant when considering transmission networks where losses can commonly be ignored. The use of the SOCP relaxation is of interest when this relaxation is exact [FL13] (typically on radial distribution networks). Since solving OPF_{AC} can be time-consuming, the number of times this problem is solved should be kept small and if possible, it should be solved on instances of relatively small scale when it is crucial to keep the solve time short.

1.5 Contributions

This dissertation gathers previous and ongoing research on different topics related to the broad theme of transmission and distribution coordination in electric power systems.

Chapter 2 suggests an iterative Gauss-Newton algorithm for solving the static version of AC-OPF and a data-driven approach aiming at identifying critical scenarios to solve a stochastic version of AC-OPF. The content of this chapter relies on the following works:

- Ilyès Mezghani, Quoc Tran-Dinh, Ion Necoara and Anthony Papavasil-

iou. A Penalty Method Based on a Gauss-Newton Scheme for AC-OPF. *2021 IEEE Madrid PowerTech*, IEEE, 2021. (Extended version on arXiv: `https://arxiv.org/abs/1905.08588`)

- Ilyès Mezghani, Sidhant Misra and Deepjyoti Deka. Stochastic AC Optimal Power Flow: A Data-Driven Approach. *Electric Power Systems Research*, 189:106567, 2020.

Chapter 3 disseminates the preliminary modeling and ideas of transmission and distribution coordination schemes proposed in the scope of SmartNet [MRS+17]. This study lead to the following publication:

- Anthony Papavasiliou and Ilyès Mezghani. Coordination Schemes for the Integration of Transmission and Distribution System Operations. *2018 Power Systems Computation Conference (PSCC)*, IEEE, 2018.

Chapter 4 implements a market-clearing tool for integrated transmission and distribution systems. This is ongoing research which lead to one conference paper and a submitted journal paper:

- Ilyès Mezghani and Anthony Papavasiliou. A Mixed Integer Second Order Cone Program for Transmission-Distribution System Co-Optimization. *2019 IEEE Milan PowerTech*, IEEE, 2019.

- *Submitted to IEEE Transactions on Power Systems.* Ilyès Mezghani, Nicolas Stevens and Anthony Papavasiliou. Hierarchical Coordination of Transmission and Distribution System Operations. 2021.

Chapter 5 aims at assessing the strengths and weaknesses of certain coordination schemes from a game-theoretical point of view. This work lead to one journal paper and one conference paper:

- Hélène Le Cadre, Ilyès Mezghani, and Anthony Papavasiliou. A Game-Theoretic Analysis of Transmission-Distribution System Operator Coordination. *European Journal of Operational Research*, 274.1:317-339, 2019.

- Ilyès Mezghani, Anthony Papavasiliou, and Hélène Le Cadre. A Generalized Nash Equilibrium Analysis of Electric Power Transmission-Distribution Coordination. *Proceedings of the Ninth International Conference on Future Energy Systems*, 2018.

1.A Optimal Power Flow Nomenclature

This section provides the main OPF nomenclature introduced in section 1.4 and used throughout the dissertation.

Sets

$\mathcal{B}$	Set of buses
$\mathcal{TB}$	Set of transmission buses
$\mathcal{DB}$	Set of distribution buses
$\mathcal{PV}$	Set of PV buses
$\mathcal{PQ}$	Set of PQ buses
$\{Slack\}$	Slack bus (singleton)
$\mathcal{L}, \mathcal{L}^R$	Set of lines, set of reversed lines
$\mathcal{TL}$	Set of transmission lines
$\mathcal{DL}$	Set of distribution lines
$\{(i', j')\}$	Interconnection line (singleton)
$\mathcal{G}$	Set of generators (or BSPs)
$\mathcal{G}_i$	Set of generators connected $i \in \mathcal{B}$
$\mathcal{TG}$	Set of transmission generators
$\mathcal{DG}$	Set of distribution generators
$\mathcal{T}$	Set of time periods
$\mathcal{OC}$	Set of operational constraints
$\mathcal{OC}^{form}$	Set of operational constraints in formulation $form = \{\text{AC-T}, \text{DC}, \text{Q}, \text{SOC}\}$
$\mathcal{IOC}$	Set of operational constraints at the interconnection
$\mathcal{GC}_g$	Set of generation constraints for generator $g \in \mathcal{G}$
$\mathcal{GC}_g^{\mathbf{DC}}$	Set of generation constraints in the DC approximation for generator $g \in \mathcal{G}$

Parameters

Y	Admittance matrix defined for every line $(i, j) \in \mathcal{L}$
G_{ij}	Conductance of line $(i, j) \in \mathcal{L}$
B_{ij}	Susceptance of line $(i, j) \in \mathcal{L}$
D_i^p	Real power demand at bus $i \in \mathcal{B}$
D_i^q	Reactive power demand at bus $i \in \mathcal{B}$
$\underline{v}_i / \overline{v}_i$	Lower/Upper voltage limit at bus $i \in \mathcal{B}$
$\underline{f}_l^p / \overline{f}_l^p$	Lower/Upper real power limit on line $l \in \mathcal{L}$
$\underline{f}_l^q / \overline{f}_l^q$	Lower/Upper reactive power limit on line $l \in \mathcal{L}$
S_l	Apparent power limit on line $l \in \mathcal{L}$
$\underline{p}_g / \overline{p}_g$	Lower/Upper real power capacity of generator $g \in \mathcal{G}$
$\underline{q}_g / \overline{q}_g$	Lower/Upper reactive power capacity of generator $g \in \mathcal{G}$
$\overline{pq}_g$	Apparent power capacity of generator $l \in \mathcal{G}$

Variables

P_i	Real power injection at bus $i \in \mathcal{B}$
Q_i	Reactive power injection at bus $i \in \mathcal{B}$
p_g	Real power generation at generator $g \in \mathcal{G}$
q_g	Reactive power generation at generator $g \in \mathcal{G}$
v_i	Voltage magnitude at bus $i \in \mathcal{B}$
θ_i	Voltage angle at bus $i \in \mathcal{B}$

f_l^p	Real power flow on line $l \in \mathcal{L} \cup \mathcal{L}^R$
f_l^q	Reactive power flow on line $l \in \mathcal{L} \cup \mathcal{L}^R$
c_{ij}	'Cosine' variable introduced in the quadratic reformulation $c_{ii} = v_i^2, c_{ij} = v_i v_j \cos(\theta_i - \theta_j),\ (i,j) \in \mathcal{L}$
s_{ij}	'Sine' variable introduced in the quadratic reformulation $s_{ii} = 0, s_{ij} = -v_i v_j \sin(\theta_i - \theta_j),\ (i,j) \in \mathcal{L}$
y_g	Binary variables associated with power generation p_g, $g \in \mathcal{G}$
λ_i	Locational Marginal Price (LMP) associated with power balance constraints at bus $i \in \mathcal{B}$
λ_i^p	Real Locational Marginal Price (LMP) associated with real power balance constraints at bus $i \in \mathcal{B}$
λ_i^q	Reactive Locational Marginal Price (LMP) associated with reactive power balance constraints at bus $i \in \mathcal{B}$
x	Vector of primal variables
x^T/x^D	Vector of primal transmission/distribution variables

Functions

$C_g(.)$	Cost function of generator $g \in \mathcal{G}$
$\delta(.)$	Function returning the set of neighbors of a bus $\delta(i) = \{j \mid (i,j) \in \mathcal{L} \cup \mathcal{L}^R\}$
F_i	Concise notation function for power balance constraints
F_i^{form}	Concise notation function for power balance constraints in formulation $form = \{\text{AC-T}, \text{DC}, \text{Q}, \text{SOC}\}$

We extend the use of subscript t to designate sets, parameters, variables or functions needed to be specified at a certain time-step t, specifically in the context of multi-period OPF. We also designate vectors of variables using this format: $z_{\mathcal{Z}}$, where z is a certain type of variable and $\mathcal{Z}$ a subset of the indices on which the variable is defined. In the same way, z_t would designate the variable z for the time-period t.

Chapter 2

Algorithms for Deterministic and Stochastic AC-OPF

2.1 Introduction

In this chapter, we aim at tackling two variants of the classical AC optimal power flow problem. First (section 2.2), we propose a novel approach for solving AC-OPF. This iterative scheme is based on Gauss-Newton theory and numerical results on several state-of-the art instances demonstrate competitive performances compared to an interior point solver, IPOPT. In a second part (section 2.3), we discuss a stochastic version of AC-OPF where the demand is uncertain. In this setting, we perform a careful critical scenario identification in order to solve a simplified version of the original problem. The method reveals robust performance on three small test cases as well as one larger instance.

2.2 A Gauss-Newton Algorithm for Solving AC-OPF

2.2.1 Related Work and Contributions

In recent years, there has been a great body of literature that has focused on convex relaxations of the AC-OPF problem, including semidefinite programming relaxations [LL11], conic relaxations [Jab08, KDS16], and quadratic relaxations [CHVH15]. These works have established conditions under which these relaxations are exact, and understanding cases in which this is not so

[MHLD13]. Instead, our interest in the present work is to tackle directly this problem as a non-convex optimization problem with non-linear equality constraints. Such a formulation is sufficiently general to produce physically implementable solutions in the context of realistic system operations.

The AC-OPF problem is usually formulated as a non-convex optimization problem. It is well known that optimization problems with non-convex constraints are difficult to solve. Classical techniques such as interior-point, augmented Lagrangian, penalty, Gauss-Newton, and sequential quadratic programming methods can only aim at finding a stationary point, which is a candidate for a local minimum [NW06]. For an iterative method to identify a stationary point that is a local minimum, but not a saddle-point, more sophisticated techniques are required, such as cubic regularization [NP06] or random noise gradient [DPG$^+$14]. However, these methods are often very difficult to implement and inefficient in large-scale problems with non-convex constraints [NP06, DPG$^+$14]. One of the most efficient and well-established non-linear solvers for finding stationary points is IPOPT [WB06], which relies on a primal-dual interior-point method combined with other advanced techniques. We emphasize that this classical method is only guaranteed to converge to a stationary point, and often requires a strategy such as line-search, filter, or trust-region to achieve global convergence under certain restrictive assumptions. Moreover, each iteration of IPOPT requires solving a non-convex subproblem via linearization combined with a line-search or filter strategy.

Recently there has been a revived interest in the design and analysis of algorithms for solving optimization problems involving non-convex constraints, in particular in engineering and machine learning [BDL19, BP16, CGT14, CRS18, TDGMD11]. The main trend is in solving large-scale problems by exploiting special structures/properties of the problem model and data towards the design of simple schemes (e.g., solving a tractable convex subproblem at each iteration), while producing reasonable approximate solutions efficiently [DP19, LW16, Nes07].

Following this trend, we will focus on Gauss-Newton (GN) methods for solving non-convex optimization problems. The idea of the GN method studied in this work was proposed in [LW16] for minimizing a compositional model of the form $\phi(F(x))$, where F is convex and ϕ is possibly nonsmooth. The main idea is to replace the continuous least-squares function $\phi(\cdot) = \frac{1}{2}\|\cdot\|_2^2$ in these methods by a given convex and Lipschitz continuous function $\phi(\cdot)$ (but possibly nonsmooth). Nesterov contributed a thorough investigation on convergence guarantees of this method in [Nes07] when ϕ is a given norm. This was extended to a more general model that can cover exact penalty methods in a technical report [TDD11].

Our approach and contributions In this work, we consider an AC-OPF problem over large-scale networks. We are interested in an approach that

can tackle general meshed networks. We present how this problem can be posed in the framework of non-convex optimization with a particular structure on the constraints. Based on this structure we devise a provably convergent Gauss-Newton (GN)-type algorithm for solving this non-convex problem. Our algorithm converges globally to a stationary point of the problem from any starting point. In addition, it is also different from standard GN methods in the literature due to the use of a non-smooth penalty instead of a classical quadratic penalty term. Hence, we refer to this algorithm as a *global and robust* GN scheme. The main idea of our method is to keep the convex substructure of the original problem unchanged and to convexify the non-convex part by exploiting penalty theory and the GN framework. Hence, in contrast to IPOPT, each iteration of our algorithm requires solving a convex subproblem, which can efficiently be solved by many existing convex solvers.

The main contributions of the work are the following:

(i) We consider a quadratic reformulation of the AC-OPF problem, as in [Jab08,ER99], and propose an exact penalty reformulation of the problem in order to handle the non-convex equality constraints and a novel global and robust GN algorithm for solving the corresponding problem.

(ii) For our optimization algorithm, we prove that its iterate sequence converges globally (i.e. from any starting point) to a stationary point of the underlying problem. We also estimate its best-known global sublinear convergence rate.

(iii) We show that the newly developed algorithm can be implemented efficiently on AC-OPF problems and test it on several numerical examples from the well-known MATPOWER test cases [ZMST10, EDA18]. We observe competitive performance to the well-established and widely-used IPOPT solver.

Our algorithm is simple to implement and can be incorporated flexibly with any available convex sub-solver that supports a warm start strategy in order to gain efficiency.

Content The first part of the chapter is organized as follows. In Section 2.2.2, we introduce our Gauss-Newton algorithm. In Section 2.2.3, we present the AC-OPF problem, its quadratic reformulation and the application of GN to AC-OPF, then illustrate through simulations the performance of our algorithm in Section 2.2.4. Finally, Section 2.2.5 concludes the first part of the chapter.

2.2.2 A Gauss-Newton Algorithm for Non-Convex Optimization

In this section, we rely on the following non-convex optimization problem:

$$\min_{\boldsymbol{x}\in\mathbb{R}^d} f(\boldsymbol{x}) \;\text{ s.t. }\; \Psi(\boldsymbol{x})=0,\; \boldsymbol{x}\in\mathcal{X}. \tag{2.1}$$

We present the main assumptions for (2.1), propose an exact penalty reformulation, and solve it using a Gauss-Newton-type algorithm. We further characterize the global and local convergence rates of our algorithm.

2.2.2.1 Exact Penalty Approach for Non-Convex Programming

For the non-convex optimization problem (2.1), we assume that the objective function f is convex and differentiable and $\mathcal{X}$ is a compact convex set. Note that our method developed in the sequel can also be extended to non-smooth convex function f or smooth non-convex function f whose gradient is Lipschitz continuous, but we make this assumption for simplicity of presentation. Furthermore, the non-convexity enters into the optimization problem through the non-linear equality constraints $\Psi(\boldsymbol{x})=0$ defined by $\Psi:\mathbb{R}^d\to\mathbb{R}^n$. We assume that Ψ is differentiable and its Jacobian Ψ' is Lipschitz continuous, i.e. there exists $L_\Psi>0$ such that:

$$\|\Psi'(\boldsymbol{x})-\Psi'(\hat{\boldsymbol{x}})\|\le L_\Psi\|\boldsymbol{x}-\hat{\boldsymbol{x}}\| \quad \forall \boldsymbol{x},\hat{\boldsymbol{x}}\in\mathcal{X},$$

where $\|\cdot\|$ is the ℓ_2-norm. Further, let $\mathcal{N}_\mathcal{X}$ denote the normal cone of the convex set $\mathcal{X}$:

$$\mathcal{N}_\mathcal{X}(\boldsymbol{x}):=\begin{cases}\left\{\boldsymbol{w}\in\mathbb{R}^d \mid \boldsymbol{w}^\top(\boldsymbol{y}-\boldsymbol{x})\ge 0, \forall \boldsymbol{y}\in\mathcal{X}\right\}, \text{if } \boldsymbol{x}\in\mathcal{X}\\ \varnothing, \quad \text{otherwise.}\end{cases}$$

Since problem (2.1) is non-convex, our goal is to search for a stationary point of this optimization problem that is a candidate for a local optimum in the following sense.

Definition 2.2.1 ([NW06], Theorem 12.9)**.** *A point $(\boldsymbol{x}^*,\boldsymbol{y}^*)$ is said to be a KKT point of* (2.1) *if it satisfies the following conditions:*

$$-\nabla f(\boldsymbol{x}^*)-\Psi'(\boldsymbol{x}^*)\boldsymbol{y}^*\in\mathcal{N}_\mathcal{X}(\boldsymbol{x}^*),\; \boldsymbol{x}^*\in\mathcal{X},\; \Psi(\boldsymbol{x}^*)=0.$$

Here, $\boldsymbol{x}^*$ is called a stationary point of (2.1)*, and $\boldsymbol{y}^*$ is the corresponding multiplier. Let $\mathcal{S}^*$ denote the set of these stationary points.*

Since $\mathcal{X}$ is compact, and Ψ and f are continuous, by the well-known Weierstrass theorem, we have:

Proposition 2.2.1. *If $\mathcal{X}\cap\{\boldsymbol{x}\mid\Psi(\boldsymbol{x})=0\}\neq\varnothing$, then* (2.1) *has global optimal solutions.*

2.2.2.2 Exact Penalized Formulation

Associated with (2.1), we consider its exact penalty form [NW06, Chapt. 17.3]:

$$\min_{\boldsymbol{x}\in\mathcal{X}} \Big\{ F(\boldsymbol{x}) := f(\boldsymbol{x}) + \beta|\Psi(\boldsymbol{x})|_1 \Big\}, \tag{2.2}$$

where $\beta > 0$ is a penalty parameter, and $|\cdot|_1$ is the ℓ_1-norm. Two reasons for choosing an exact (non-smooth) penalty are as follows. First, for a certain finite choice of the parameter β, a single minimization in x of (2.2) can yield an exact solution of the original problem (2.1). Second, it does not square the condition number of Ψ as in the case of quadratic penalty methods, thus making our algorithm presented below more robust to ill-conditioning of the non-convex constraints. Now, we summarize the relationship between stationary points of (2.1) and of its penalty form (2.2). For this, let us define the directional derivative:

$$DF(\boldsymbol{x}^*)[\boldsymbol{d}] := \nabla f(\boldsymbol{x}^*)^\top \boldsymbol{d} + \beta\xi(\boldsymbol{x}^*)^\top \Psi'(\boldsymbol{x}^*)^\top \boldsymbol{d}, \tag{2.3}$$

where $\xi(\boldsymbol{x}^*) \in \partial|\Psi(\boldsymbol{x}^*)|_1$ is one subgradient of $|\cdot|_1$ at $\Psi(\boldsymbol{x}^*)$[1], and $\partial|\cdot|_1$ denotes the subdifferential of $|\cdot|_1$, see [Nes07]. Recall that the necessary optimality condition of (2.2) is

$$0 \in \nabla f(\boldsymbol{x}^*) + \beta\Psi'(\boldsymbol{x}^*)\partial|\Psi(\boldsymbol{x}^*)|_1 + \mathcal{N}_{\mathcal{X}}(\boldsymbol{x}^*).$$

Then, this condition can be expressed equivalently as

$$DF(\boldsymbol{x}^*)[\boldsymbol{d}] \geq 0, \quad \forall \boldsymbol{d} \in \mathcal{F}_{\mathcal{X}}(\boldsymbol{x}^*), \tag{2.4}$$

where $\mathcal{F}_{\mathcal{X}}(\boldsymbol{x})$ is the set of feasible directions to $\mathcal{X}$ at $\boldsymbol{x}$:

$$\mathcal{F}_{\mathcal{X}}(\boldsymbol{x}) := \left\{ \boldsymbol{d} \in \mathbb{R}^d \mid \boldsymbol{d} = t(\boldsymbol{y} - \boldsymbol{x}),\ \forall \boldsymbol{y} \in \mathcal{X},\ t \geq 0 \right\}.$$

Any point $\boldsymbol{x}^*$ satisfying (2.4) is called a stationary point of the penalized problem (2.2). Stationary points are candidates for local minima, local maxima, and saddle-points. If, in addition, $\boldsymbol{x}^*$ is feasible to (2.1), then we say that $\boldsymbol{x}^*$ is a feasible stationary point. Otherwise, we say that $\boldsymbol{x}^*$ is an infeasible stationary point. Proposition 2.2.2 shows the relation between (2.1) and (2.2).

Proposition 2.2.2 ([NW06], (Theorem 17.4.)). *Suppose that $\boldsymbol{x}^*$ is a feasible stationary point of* (2.2) *for β sufficiently large. Then, $\boldsymbol{x}^*$ is also stationary point of the original problem* (2.1).

Proposition 2.2.2 requires $\boldsymbol{x}^*$ to be feasible for (2.1). When the feasible set $\mathcal{X} \cap \{\boldsymbol{x} \mid \Psi(\boldsymbol{x}) = 0\} \neq \varnothing$ of (2.1) is nonempty and bounded, according to [DP94, Proposition 2], if (2.1) satisfies the extended Mangarasian-Fromovitz constraint

[1] Any subgradient of the subdifferential would provide the same theoretical guarantee.

qualification condition (see [DP94, Proposition 2] for concrete definition), then there exists $\beta_* > 0$ such that for any $\beta > \beta_*$, every global or local solution of the penalized problem (2.2) is also a global or local optimal solution of (2.1), respectively. By [DP94, Proposition 3], β needs to be chosen such that $\beta > \beta_* := \|\boldsymbol{y}^*\|_\infty$, where $\boldsymbol{y}^*$ is any optimal Lagrange multiplier of (2.1). The intuition of the claim 'for β sufficiently large' could be explained as follows: assuming $\beta = +\infty$, the only way to reach a minimizer of (2.2) is to have $\Psi(\boldsymbol{x}) = 0$ which is then a feasible point of (2.1). Now that $\beta|\Psi(\boldsymbol{x})|_1 = 0$, minimizing the objective of (2.2) leads to minimizing $f(x)$, which is why it is equivalent to solving (2.1).

2.2.2.3 Global Gauss-Newton Method

Our GN method aims at solving the penalized problem (2.2) using the following convex subproblem:

$$\min_{\boldsymbol{x}\in\mathcal{X}} \Big\{ \mathcal{Q}_L(\boldsymbol{x};\boldsymbol{x}^k) := f(\boldsymbol{x}) + \beta|\Psi(\boldsymbol{x}^k) + \Psi'(\boldsymbol{x}^k)(\boldsymbol{x}-\boldsymbol{x}^k)|_1 + \tfrac{L}{2}\|\boldsymbol{x}-\boldsymbol{x}^k\|^2 \Big\} \quad (2.5)$$

where $\boldsymbol{x}^k$ is a given point in $\mathcal{X}$ for linearization, $\Psi'(\cdot)$ is the Jacobian of Ψ, and $L > 0$ is a regularization parameter.

Note that our subproblem (2.5) differs from those used in classical penalty methods [NW06], since we linearize the constraints and we also add a regularization term. Thus, the objective function of (2.5) is strongly convex. Hence, if $\mathcal{X}$ is nonempty and even if the problem is non-differentiable, this problem admits a unique optimal solution, and can be solved efficiently by several convex methods and solvers. For instance, alternating direction methods of multipliers (ADMM) [BPC+11] and primal-dual schemes [CP11] can be efficient for solving (2.5). Note that the convergence guarantees of ADMM and primal-dual schemes often depends on the distance between the initial point $\boldsymbol{x}^{k,0}$ of the algorithm and the exact optimal solution of $\bar{\boldsymbol{x}}^{k+1}$ of (2.5), see, e.g, [CP11, Theorem 2]. Hence, if we warm-start $\boldsymbol{x}^{k,0}$ at the previous approximate solution $\boldsymbol{x}^k$ obtained at the $(k-1)$-th iteration, then the distance $\|\boldsymbol{x}^0 - \bar{\boldsymbol{x}}^{k+1}\|$ is small. This allows the algorithm to converge faster to a desired approximate solution $\boldsymbol{x}^{k+1}$ of (2.5) at the k-th iteration.

Let us define:

$$\boldsymbol{V}_L(\boldsymbol{x}^k) := \operatorname*{argmin}_{x\in\mathcal{X}} \Big\{ \mathcal{Q}_L(\boldsymbol{x};\boldsymbol{x}^k) \Big\}. \quad (2.6)$$

And the following quantities:

$$\begin{aligned} &\boldsymbol{G}_L(\boldsymbol{x}^k) := L(\boldsymbol{x}^k - \boldsymbol{V}_L(\boldsymbol{x}^k)), \\ &\boldsymbol{d}_L(\boldsymbol{x}^k) := \boldsymbol{V}_L(\boldsymbol{x}^k) - \boldsymbol{x}^k, \\ &r_L(\boldsymbol{x}^k) := \|\boldsymbol{d}_L(\boldsymbol{x}^k)\|. \end{aligned} \quad (2.7)$$

$\boldsymbol{G}_L(\cdot)$ can be considered as a gradient mapping of F in (2.2) [Nes07], and $\boldsymbol{d}_L(\boldsymbol{x}^k)$ is a search direction for Algorithm 1. The necessary and sufficient

optimality condition for subproblem (2.5) becomes:

$$\left[\nabla f(\boldsymbol{V}_L(\boldsymbol{x}^k)) - \boldsymbol{G}_L(\boldsymbol{x}^k) + \beta\Psi'(\boldsymbol{x}^k)\xi(\boldsymbol{x}^k)\right]^\top (\hat{\boldsymbol{x}} - \boldsymbol{V}_L(\boldsymbol{x}^k)) \geq 0, \quad \forall \hat{\boldsymbol{x}} \in \mathcal{X}, \tag{2.8}$$

where $\xi(\boldsymbol{x}^k) \in \partial|\Psi(\boldsymbol{x}^k) + \Psi'(\boldsymbol{x}^k)(\boldsymbol{V}_L(\boldsymbol{x}^k) - \boldsymbol{x}^k)|_1$. Now, using the subproblem (2.5) as a main component, we describe our GN scheme in Algorithm 1.

Algorithm 1 *The Basic Gauss-Newton Algorithm*

1: **Initialization:** Choose $\boldsymbol{x}^0 \in \mathcal{X}$ and a penalty parameter $\beta > 0$ sufficiently large (ideally, $\beta > \|\boldsymbol{y}^*\|_\infty$, where $\boldsymbol{y}^*$ is a dual optimal solution of (2.1)).
2: Choose a lower bound $L_{\min} \in (0, \beta L_\Psi]$.
3: **For** $k := 0$ **to** $k_{\max}$ **perform**
4: Find $L_k \in [L_{\min}, \beta L_\Psi]$ such that $F(\boldsymbol{V}_{L_k}(\boldsymbol{x}^k)) \leq \mathcal{Q}_{L_k}(\boldsymbol{V}_{L_k}(\boldsymbol{x}^k); \boldsymbol{x}^k)$.
5: Update $\boldsymbol{x}^{k+1} := \boldsymbol{V}_{L_k}(\boldsymbol{x}^k)$.
6: Update β if necessary.
7: **End for**

The global and local convergence analysis of Algorithm 1 is discussed in Appendix 2.A.

The main step of Algorithm 1 is the solution of the convex subproblem (2.5) at Step 4. As mentioned, this problem is strongly convex, and can be solved by several methods that converge linearly. If we choose $L_k \equiv L \geq \beta L_\Psi$, then we do not need to perform a line-search on L at Step 4, and only need to solve (2.5) once per iteration. However, L_Ψ may not be known or if it is known, the global upper bound βL_Ψ may be too conservative, i.e. it does not take into account the local structures of non-linear functions in (2.2). Therefore, following the algorithm in [Nes07], we propose performing a line-search in order to find an appropriate L_k. If we perform a line-search by doubling L_k at each step starting from $L_{\min}$, (i.e., $L_k \to 2L_k$), then after i_k line-search steps, we have $L_k = 2^{i_k} L_{\min}$, and the number of line-search iterations i_k is at most $\lfloor \log_2(\beta L_\Psi / L_{\min}) \rfloor + 1$. A detailed discussion on tuning β and L are provided in Appendix 2.B.

2.2.2.4 The Basic Gauss-Newton Algorithm versus IPOPT

We summarize in Table 2.1 the main differences between the GN scheme suggested in this work and the classical interior point solver, IPOPT. To make the comparison clearer, IPOPT is usually presented by writing an NLP in the following format:

$$\min_{\boldsymbol{x} \in \mathbb{R}^d} f(\boldsymbol{x}) \;\; \text{s.t.} \;\; \Lambda(\boldsymbol{x}) = 0, \;\; \boldsymbol{x} \geq 0. \tag{2.9}$$

where Λ is twice continuously differentiable.

Table 2.1: Main differences between the proposed GN algorithm and IPOPT.

Aspect	GN Algorithm	IPOPT
Assumptions on the non-linear constraints	Ψ differentiable and Jacobian Ψ' Lipschitz continuous.	Λ twice continuously differentiable.
Penalization	It enters the scope of 'penalty methods' and relies on penalizing the non-linear constraints using the ℓ_1-norm: $\min_{\boldsymbol{x}\in\mathbb{R}^d} \quad f(\boldsymbol{x}) + \beta\lvert\Psi(x)\rvert_1$ s.t. $\quad \boldsymbol{x} \in \mathcal{X}$	It enters the scope of 'barrier methods' and relies on a logarithmic penalization of the variables (μ is a parameter[2]): $\min_{\boldsymbol{x}\in\mathbb{R}^d} \quad f(\boldsymbol{x}) - \mu\sum_{i=1}^{d}\ln(x_i)$ s.t. $\quad \Lambda(\boldsymbol{x}) = 0$
Iteration	Consists of solving QCP (2.5). This is efficiently solved by the homogeneous and self-dual algorithm which is implemented in commercial solvers like Mosek.	Consists of solving a barrier problem, i.e. a system of non-linear equations. Note that a lot of strong features of IPOPT are implemented to accelerate the execution of one iteration.
Global convergence guarantee	First order convergence guarantee to a stationary point of (2.2).	Under mild assumptions [Wäc02], IPOPT converges to a stationary point of (2.9).

Our GN scheme and IPOPT are based on different approaches for solving NLPs. Our GN scheme applies to more general problems but IPOPT offers more accurate convergence guarantee (stationary point of the penalized problem (2.2) for GN as opposed to stationary point of the original nonconvex problem for IPOPT) on the problem that is discussed here, namely AC-OPF.

2.2.3 Applying the GN Algorithm to AC-OPF

In this section, we present the OPF problem and its reformulation in a form that obeys the structure presented in the previous section. We then perform numerical experiments to validate the algorithm and compare it with IPOPT.

[2] μ is updated to tend to 0 when the algorithm converges. In the GN algorithm, β is fixed to a large value and if updated, L increases within an iteration.

2.2.3.1 Problem Settings

The starting point of our proposed GN method for solving problem OPF_{AC} (problem (1.6)) is the equivalent trigonometric reformulation of AC-OPF [ER99] formulated as $\text{OPF}_{\text{AC-T}}$ (problem (1.11)).

In this reformulation, we have two non-convex equality constraints:

- Constraints (1.9a): $\Psi_q^{ij}(\boldsymbol{c}, \boldsymbol{s}) := c_{ij}^2 + s_{ij}^2 - c_{ii}c_{jj} = 0,\ \forall (i,j) \in \mathcal{L}$. We refer to them as *quadratic* constraints.

- Constraints (1.9b): we slightly rewrite these constraints and cast them as follows: $\Psi_t^{ij}(\boldsymbol{c}, \boldsymbol{s}, \boldsymbol{\theta}) := \sin(\theta_i - \theta_j)c_{ij} + \cos(\theta_i - \theta_j)s_{ij} = 0,\ \forall (i,j) \in \mathcal{L}$. We refer to them as *trigonometric* constraints.

Since Ψ is the collection of (1.9a) and (1.9b), we show in the next lemma that it is differentiable and that its Jacobian is Lipschitz continuous.

Lemma 2.2.1. *For the AC-OPF problem, Ψ defined by* (1.9a)–(1.9b)*, is smooth, and its Jacobian Ψ' is Lipschitz continuous with a Lipschitz constant L_Ψ, i.e. $\|\Psi'(\boldsymbol{x}) - \Psi'(\hat{\boldsymbol{x}})\| \leq L_\Psi \|\boldsymbol{x} - \hat{\boldsymbol{x}}\|$ for all $\boldsymbol{x}, \hat{\boldsymbol{x}} \in \mathcal{X}$, where*

$$L_\Psi := \max\left\{2, \left(1 + 2\max\{\overline{v}_i^2 \mid i \in \mathcal{B}\}\right)^{1/2}\right\} < +\infty. \tag{2.10}$$

Proof. From the definition of Ψ, it consists of two parts: quadratic forms in (c_{ij}, s_{ij}) and trigonometric and linear forms in $\theta_{ij} := \theta_i - \theta_j$ and (c_{ij}, s_{ij}), respectively. We can write it as $\Psi = [\Psi^q, \Psi^t]$. Each function in Ψ^q has the form $c_{ij}^2 + s_{ij}^2 - c_{ii}c_{jj}$, as shown by (1.9a), and each function in Ψ^t has the form $\sin(\theta_{ij})c_{ij} + \cos(\theta_{ij})s_{ij}$, as shown in (1.9b). We can show that the second derivative of each component of Ψ^q w.r.t. $(c_{ii}, c_{jj}, c_{ij}, s_{ij})$ and of Ψ^t w.r.t. $(c_{ij}, s_{ij}, \theta_{ij})$, respectively is

$$\nabla^2\Psi^q(\boldsymbol{x}) = \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \quad \text{and}$$

$$\nabla^2\Psi^t(\boldsymbol{x}) = \begin{bmatrix} 0 & 0 & \cos(\theta_{ij}) \\ 0 & 0 & -\sin(\theta_{ij}) \\ \cos(\theta_{ij}) & -\sin(\theta_{ij}) & -c_{ij}\sin(\theta_{ij}) - s_{ij}\cos(\theta_{ij}) \end{bmatrix}.$$

The second derivative $\nabla^2\Psi^q(\boldsymbol{x})$ is constant. Hence, the maximum eigenvalue of $\nabla^2\Psi^q(\boldsymbol{x})$ is $\lambda_{\max}(\nabla^2\Psi^q(\boldsymbol{x})) = 2$. For any $\boldsymbol{u} := [u_1, u_2, u_3] \in \mathbb{R}^3$, and $|.|$

denoting the absolute value, we can easily estimate that

$$\begin{aligned}\boldsymbol{u}^\top \nabla^2 \Psi^t(\boldsymbol{x})\boldsymbol{u} &= u_1 u_3[-\sin(\theta_{ij}) + \cos(\theta_{ij})] + u_2 u_3[\cos(\theta_{ij}) \\ &\quad - \sin(\theta_{ij})] + [-c_{ij}\sin(\theta_{ij}) - s_{ij}\cos(\theta_{ij})]u_3^2 \\ &\leq \tfrac{1}{2}(u_1^2 + u_3^2) + \tfrac{1}{2}(u_2^2 + u_3^2) + (|c_{ij}| + |s_{ij}|)u_3^2 \\ &\leq (1 + |c_{ij}| + |s_{ij}|)(u_1^2 + u_2^2 + u_3^2) \\ &= (1 + |c_{ij}| + |s_{ij}|)\|u\|^2.\end{aligned}$$

Therefore,

$$\lambda_{\max}\left(\nabla^2 \Psi^t(\boldsymbol{x})\right) = 1 + |c_{ij}| + |s_{ij}| \leq 1 + \max\{c_i \mid i \in \mathcal{B}\} + \max\{s_i \mid i \in \mathcal{B}\}$$

And from (1.10e), we have:

$$L_\Psi := \max\left\{2, \left(1 + 2\max\{\overline{v}_i^2 \mid i \in \mathcal{B}\}\right)^{1/2}\right\} < +\infty,$$

which is (2.10). □

Moreover, AC-OPF is written in the format of (2.1), where:

- $\boldsymbol{x} := (\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{c}, \boldsymbol{s}, \boldsymbol{\theta}, \boldsymbol{f^p}, \boldsymbol{f^q})$, $f(\boldsymbol{x}) := \sum_{g \in \mathcal{G}} C_g(p_g)$,
- $\Psi(\boldsymbol{x}) := (\Psi_q^{ij}(\boldsymbol{c}, \boldsymbol{s}), \Psi_t^{ij}(\boldsymbol{c}, \boldsymbol{s}, \boldsymbol{\theta}))$
- and $\mathcal{X} := \{\boldsymbol{x}$ satisfies (1.10c), (1.10d), (1.10e), (1.5g), (1.6d)$\}$ is a convex set gathering linear and second-order cone constraints.

We can now derive AC-OPF subproblems for GN. By using the following equivalence,

$$\begin{array}{lll} \min\limits_{z} & |z|_1 & \\ s.t. & z \in \mathcal{Z}\ (\subseteq \mathbb{R}^d) & \end{array} \quad \Longleftrightarrow \quad \begin{array}{lll} \min\limits_{z,t} & \sum\limits_{j=1}^{d} t_j & \\ s.t. & z_j \leq t_j & \\ & -z_j \leq t_j & \\ & z \in \mathcal{Z},\ t \in \mathbb{R}^d & \end{array}$$

we cast the objective of (2.5) as convex quadratic. We typically refer to an SOCP as an optimization problem with second-order cone constraints and a linear objective. As a consequence, the GN subproblem is a Quadratic Convex Program (QCP). We are now in a position to apply GN to AC-OPF.

2.2.3.2 A Practical Implementation of the GN Algorithm for OPF

In this section, the goal is to optimize the settings of the GN method. This will allow us to derive a practical version of the GN algorithm, which we compare to IPOPT.

Stopping criteria. We terminate Algorithm 1 in two occasions, which have been validated through experimental results:

- If the maximum number of iterations $k_{\max} := 100$ is reached.
- If the quadratic and trigonometric constraints are satisfied with a tolerance of ϵ_2, where $\epsilon_2 := 1e^{-5}$. Concretely, we stop Algorithm 1 if:

$$\max(\|\Psi_q(c^k, s^k)\|_\infty, \|\Psi_t(c^k, s^k, \theta^k)\|_\infty) \leq \epsilon_2.$$

where $\|.\|_\infty$ is the infinity norm. If the difference $\|x^{k+1} - x^k\|_\infty < \epsilon_1$ ($\epsilon_1 := 1e^{-6}$ in the numerical experiment), then the last iterate might not be feasible for (1.11). We then use the run-and-inspect strategy [CSY19]: (i) *run* until one stopping criterion is reached, (ii) *inspect* if the solution obtained is feasible and if not *run* GN one more time: the last iterate becomes the starting point of GN and β is doubled.

SOCP relaxation for initialization. As mentioned previously, the quadratic formulation is also used to derive the SOCP relaxation (1.14). In this relaxation, the angles $\boldsymbol{\theta}$ are not modeled and the trigonometric constraints (1.9b) are removed. Moreover, the non-convex constraints (1.9a) are relaxed to (1.13). Solving this relaxation will provide a partial initial point $(\boldsymbol{p}^0, \boldsymbol{q}^0, \boldsymbol{c}^0, \boldsymbol{s}^0)$ (and $\boldsymbol{\theta}^0 = 0$).

Parameter tuning strategies. Algorithm 1 presented above does not require tuning the β and L parameters. In practice, tuning is crucial for improving the performance of algorithms for constrained non-convex optimization, including Algorithm 1. Several observations allow us to decrease the number of iterations that are required for convergence: (i) according to Proposition 2.2.2, large values of β ensure the equivalence between (2.1) and (2.2); (ii) quadratic constraints and trigonometric constraints scale up differently; (iii) a careful updating of L influences the number of times that subproblem (2.5) is solved. These observations guide a detailed experimental investigation concerning the choices of β and L parameters, which is discussed in Appendix 2.B.

Acceleration through warmstart. We observe that the subproblems (2.5) are based on the same formulation, and only differ by slight changes of certain parameters along the iterations. This motivates us to warm-start

the subproblem (2.5) with a previous primal-dual iterate. In other words, we initialize the solver for solving the subproblem (2.5) at the k-th iteration at the solution $\boldsymbol{x}_{k-1}$ obtained from the previous iteration $k-1$. Note that certain solvers allow for initializing the dual values, so we also initialize the dual values to the previous iterate.Warm-starting is indeed a key step in iterative methods, including our GN scheme, and will be further analyzed in Section 2.2.4.2.

2.2.4 Numerical Experiments

In order to validate the proposed GN algorithm, our numerical experiments are conducted in 2 steps: first, we launch simulations on several test cases of a classical library (MATPOWER) and compare the GN algorithm with a state-of-the-art non-convex solver (IPOPT); second, we show the potential benefit of warm-start for our approach.

2.2.4.1 Illustration on MATPOWER Instances

We use the MATPOWER [ZMST10] library to have access to a wide range of AC-OPF test systems that have been investigated in the literature. We test our approach on instances whose size ranges between 1,354 and 25,000 nodes (`1354pegase` has 11,192 variables and 27,911 constraints while `ACTIVS25k` has 186,021 variables and 431,222 constraints).

We benchmark our approach against IPOPT, a non-linear solver based on the interior-point method. IPOPT is considered as the state-of-the-art solver for this type of problem and serves as a benchmark in [CBS+18b, EDA18, KFS18]. We make use of `PowerModels.jl` [CBS+18b], a Julia package that can be used to solve AC-OPF instances of different libraries with different formulations. In order to make a fair comparison, we initialize GN and IPOPT using the SOCP solution.

The results of our analysis are presented in Table 2.2. In Table 2.2, for each test case (first column), we report the objective value and the execution time (in seconds). For GN, we also report the number of iterations. For IPOPT, we report the solve time of the solver warm-started with the SOCP solution and without warm-start (Basic). The last column provides the gap between the GN solution and the IPOPT solution.

The first notable observation is that GN finds a stationary point (i.e. feasible) of the original AC-OPF problem for all 23 test cases. The stationary point obtained by GN attains the same objective function value as the one returned by IPOPT for most instances (a difference of 0.05% in the objective value may be attributed to numerical precision) and the proposed method outperforms IPOPT in some instances (e.g. `2737sop`).

Our experiments further demonstrate that the GN method consistently re-

Table 2.2: Comparison of the GN algorithm against IPOPT

	Gauss-Newton			IPOPT			
Test Case	# It	Objective	Time	Objective	Time (SOCP)	Time (Base)	Gap
1354pegase	13	$7.407e^4$	15.1 s	$7.407e^4$	**6.00** s	6.75 s	0.0 %
1888rte	14	$5.981e^4$	**25.9** s	**5.980e4**	59.0 s	53.7 s	0.0 %
1951rte	4	$8.174e^4$	**6.94** s	$8.174e^4$	7.67 s	22.0 s	0.0 %
ACTIVSg2000	5	$1.229e^6$	**8.39** s	$1.229e^6$	14.8 s	24.3 s	0.0 %
2383wp	19	$1.868e^6$	51.0 s	$1.868e^6$	21.6 s	**19.2** s	0.0 %
2736sp	4	**1.307e6**	**10.9** s	$1.308e^6$	13.3 s	14.4 s	-0.1 %
2737sop	3	**7.767e5**	**7.77** s	$7.778e^5$	9.98 s	12.2 s	-0.1 %
2746wop	17	$1.208e^6$	38.1 s	$1.208e^6$	**12.8** s	12.9 s	0.0 %
2746wp	3	**1.631e6**	**6.89** s	$1.632e^6$	16.0 s	14.2 s	-0.1 %
2848rte	17	$5.303e^4$	**52.0** s	**5.302e4**	74.3 s	142 s	0.0 %
2868rte	4	$7.980e^4$	**9.11** s	**7.979e4**	29.3 s	63.1 s	0.0 %
2869pegase	12	$1.340e^5$	41.9 s	$1.340e^5$	**15.5** s	35.8 s	0.0 %
3012wp	8	$2.593e^6$	25.2 s	**2.592e6**	**18.8** s	23.4 s	0.0 %
3120sp	13	**2.142e6**	41.2 s	$2.143e^6$	**19.2** s	22.2 s	0.0 %
3375wp	8	$7.413e^6$	34.7 s	**7.412e6**	**21.3** s	24.4 s	0.0 %
6468rte	19	$8.685e^4$	187 s	**8.683e4**	**103** s	139 s	0.0 %
6470rte	11	$9.835e^4$	**89.3** s	$9.835e^4$	144 s	129 s	0.0 %
6495rte	12	$1.063e^5$	180 s	$1.063e^5$	**52.4** s	121 s	0.0 %
6515rte	18	$1.098e^5$	204 s	$1.098e^5$	**86.9** s	104 s	0.0 %
9241pegase	17	$3.167e^5$	894 s	**3.159e5**	368 s	**117** s	0.3 %
ACTIVSg10k	6	$2.488e^6$	117 s	**2.486e6**	**93.4** s	355 s	0.1 %
13659pegase	19	$3.885e^5$	**137** s	**3.861e5**	685 s	710 s	0.6 %
ACTIVSg25k	16	$6.033e^6$	1,740 s	**6.018e6**	544 s	**398** s	0.3 %

quires a small number of iterations[3] (less than 20) in a wide range of instances. This is critically important to further accelerate the performance of our method if we appropriately exploit warm-start strategies and efficient solvers for the strongly convex subproblems. In terms of computational time, the performance is shared between the two approaches, and it appears like warm-starting IPOPT with the SOCP solution decreases the execution time for most of the instances but can sometimes have the opposite effects (for example 9241pegase). Nevertheless, some instances reveal limitations of the GN algorithm, compared to IPOPT (6495rte, 9241pegase and ACTIVSg25k for example): when the solution of a subproblem becomes time-consuming because of the size of the subproblem, GN might require a larger execution time. We use Gurobi for solving subproblem (2.5), because it is one of the most stable QCP solvers that are available.

[3]One iteration involves going through instructions 4, 5 and 6 of the loop 3-7 in Algorithm 1. This means that one or several optimization problems might be solved in one iteration.

Unfortunately, Gurobi (and IPMs for QCPs in general) does not support warm-start[4], which would have significantly decreased the computational time. One alternative is to use an ADMM solver that supports warm-start. However, ADMM solvers are not mature enough to test large-scale problems. Implementing an efficient subsolver is out of the scope of this work, however we are able to analyze the effect of warm start on these solvers which is the subject of the next section.

2.2.4.2 The Effect of Warm-Start Strategies

We consider using OSQP [SBG$^+$20] as an ADMM solver. Since OSQP is not as mature as commercial solvers like Gurobi and IPOPT, we only consider small test cases (`39_epri` and `118_ieee`). The results are presented in Table 2.3. From left to right, # It reports the number of GN iterations, # ADMM It reports the total number of ADMM iterations performed during the GN execution, and Time reports the sum of OSQP solve times along the iterations in seconds.

Table 2.3: Results with and without warm-start on 2 small instances using OSQP.

	No warm-start			With warm-start		
Test Case	# It	# ADMM It	Time	# It	# ADMM It	Time
`39_epri`	4	114,735	4.65 s	4	56,988	2.15 s
`118_ieee`	4	162,251	29.2 s	4	65,832	11.6 s

For both cases, we observe that warm start divides the total number of ADMM iterations as well as solve time by more than a factor of 2, and almost by a factor of 3 for `118_ieee`.

We also examine each GN iteration individually, and highlight the impact of warm-start on the number of ADMM iterations in Figure 2.1. To help with the understanding of Figure 2.1, the run of GN with 'No Warm-start' takes 100% (in blue) of the ADMM iterations and we measure the number of ADMM iterations of 'With Warm-start' relative to 'No Warm-start'. On the left graph, 2-2 represents the second subproblem that has to be solved at iteration 2 of GN due to an update of L: thanks to warm-start, 75% of ADMM iterations were saved for this particular subproblem. Note that we warm-start dual and primal variables only after iteration 1. Warm-start decreases substantially the number of ADMM iterations in two cases:

[4]Note that since there are still second order cone constraints in (2.5) (in set $\mathcal{X}$), it is not possible to use an active-set algorithm [FKP$^+$14] to exploit warm-start.

(i) When L is updated. Indeed, updating L only results in slightly changing the objective function. One expects the previous iterates to provide a good warm-start. This is confirmed in Figure 2.1, where we observe that the number of ADMM iterations is divided by at least a factor of 2 every time L is updated.

(ii) When the last iterates are computed. Intuitively, one does not expect iterates to change substantially when approaching the optimal solution. This intuition is confirmed by Figure 2.1. For the particular case of the last iterate, for `39_epri` (resp. `118_ieee`), the required number of ADMM iterations is less than 20% (resp. 30%).

This investigation suggests that, with a mature ADMM solver, warm-starting is a promising feature for improving the performance of GN on large test cases.

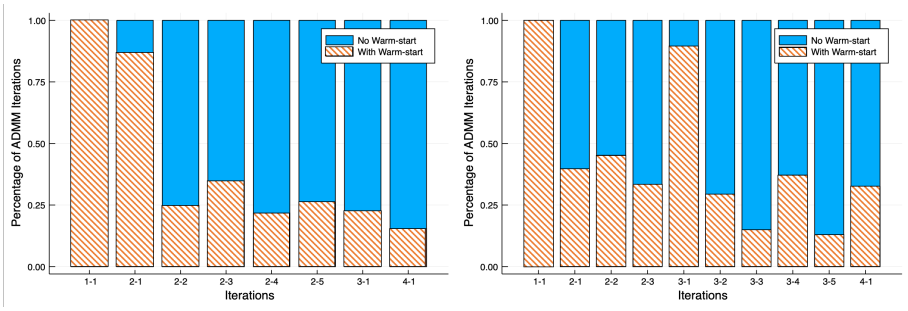


Figure 2.1: Evolution of the percentage of ADMM iterations along the iterations for `39_epri` (left) and `118_ieee` (right). 'No Warm-start' always implies 100% and the percentage of ADMM iterations 'With Warm-start' is measured relatively to 'No Warm-start'. GN iterations are shown on the x axis in an $a-b$ format: a is the actual GN iteration and b represents the bth subproblem that had to be solved at iteration a because of an update of L.

2.2.5 Conclusion

We propose a novel Gauss-Newton algorithm for solving a general class of optimization problems with non-convex constraints. We utilize an exact non-smooth penalty reformulation of the original problem and suggest an iterative scheme for solving this penalized problem which relies on the non squared Gauss-Newton method. The subproblems of our proposed scheme are strongly convex programs, which can be efficiently solved by numerous third-party convex optimization solvers.

We apply our approach to solve AC-OPF, which is a fundamental and ubiquitous problem in power systems engineering, and we propose numerous strategies for tuning our GN scheme, initializing the algorithm, and warm-starting

the resolution of the subproblems that are treated by our proposed method. We perform extensive numerical experiments on a large set of instances from the MATPOWER library, and demonstrate the competitive performance of our method to IPOPT, which is a state of the art non-linear non-convex solver. This analysis validates the theoretical analysis of our proposed GN scheme, and proves its effectiveness in practical applications.

2.3 Stochastic AC OPF: A Data-Driven Approach

2.3.1 Introduction

Modern power systems are faced with significant uncertainty in power generation and demand. This is due to the increasing integration of renewable energy resources like wind and solar, and growth of demand side participation and distributed energy resources at the sub-transmission and distribution levels. As a result, uncertainty management has become a critical component in the operational planning stage, where generators and controllable elements must be dispatched in a way that the system remains within its safety limits despite uncertain fluctuations. In the literature, the issue is addressed by considering variants of the optimal power flow problem that incorporate the effect of uncertainty. These formulations take the form of either a stochastic or robust optimization problem, where a limit on some chosen measure of risk is explicitly enforced.

All uncertainty-aware OPF formulations pose significant computational challenges, most of which can be traced back to the non-linear nature of the AC power flow equations. The two primary challenges are (i) quantifying the effect of uncertainty on the system – it is difficult to precisely express the variation of the dependent physical quantities in the system such as voltage magnitudes and line currents as a function of the uncertainty, and (ii) formulating a sufficiently compact optimization that integrates the uncertainty quantification while still being tractable. These challenges have been echoed in several recent publications on the topic [RA17, MRH+19], and several solution approaches have been proposed. Broadly, we can classify these approaches into two types (i) approximations to the AC power flow equations, and (ii) Monte Carlo methods.

Power flow approximations: These classes of approaches aim at simplifying the task of uncertainty quantification by full or partial approximations to the power flow equations. These include linear approximations of the power flow such as the DC approximation [SJA09] and a first order Taylor expansion [DBC17]. Using these approximations greatly improves tractability, in particular for risk metrics that can be expressed as a convex program. Many publications [RA17, LS17, MR18, VHM+17, MRH+19, DBS17] have attempted to incorporate the AC power flow equations. In [RA17], only a partial lineariza-

tion is considered, where the variables follow the full non-linear AC-PF while the effect of uncertainty is expressed via linearization. The resulting method is much more accurate than full linearization, but can lose fidelity when the magnitude of the uncertainty is large. More recently, an approach based on polynomial chaos expansion [MRH$^+$19] has been proposed that is highly accurate but computationally challenging. In summary, approaches based on power flow approximation trade-off accuracy for scalability – the brief review mentioned above cites methods with *high scalability - low accuracy* to *low scalability - high accuracy*.

Monte Carlo methods: These methods quantify the effect of uncertainty on the system by solving the power flow equations for a large number of realizations drawn from the uncertainty distribution. When the number of samples used is sufficiently large, Monte Carlo provides excellent accuracy. The primary challenge however lies in integrating Monte Carlo into an optimization formulation without exploding the size of the problem and the associated computational time. In this context, the most widely used approach is the so-called *scenario approach* where an extended OPF problem is formulated by incorporating a specified number of scenarios from the distribution, and robustness to each scenario is enforced via constraints. Several theoretical results [CC06, VMLA13] (primarily for convex formulations with chance constraints) provide guidelines on how many random samples should be used to achieve the desired probability of constraint violation. The main drawback of the approach is that random sampling based approaches, specifically for non-linear and non-convex optimal power flow problems, quickly result in the optimization problem becoming computationally intractable for practical cases.

Contribution: In this work, we adopt the scenario-based approach described above. However instead of including scenarios collected randomly, we use system knowledge and data-driven tools to drastically reduce the number of scenarios required to solve the problem. This keeps the primary benefits of the scenario approach – accurate uncertainty quantification, agnosticism to uncertainty distribution, etc., while significantly boosting its scalability. Our approach is an advanced iterative procedure similar to scenario generation algorithms common in problems such as power systems expansion planning [MSC$^+$17]. The algorithm iteratively adds more scenarios to the scenario-based OPF until a security criterion, assessed by a sufficient number of Monte Carlo samples, is satisfied. Note that since the assessment of scenarios *does not* involve solving the OPF, scalability is not compromised and the procedure can heavily exploit availability of parallel computing capabilities. Following the scenario assessment, what is added back to the OPF, in each iteration, is a well-chosen subset of 'modified' scenarios. To determine the 'modified' scenarios, we first develop metrics for sub-selecting a very small portion of critical scenarios based on constraint violation. Second, we use regularized linear regression to identify the directions of uncertainty that are the most adversarial

for each violated constraint. We then boost the chosen critical scenarios along the directions identified and add them back to the OPF formulation for the next iteration. We show through several case studies that this data-driven program significantly reduces the scenario size requirements over vanilla random sampling - with ~ 30 scenarios we are able to find a secure solution to the stochastic OPF for the large 1354-bus system. In summary, our contribution in this work is a suite of data-driven tools to efficiently solve the stochastic OPF problem with a scenario-based approach. The features used in our scenario selection procedure can be tuned based on historical knowledge/expertise available with an operator.

2.3.2 Problem Formulation

In this section, we provide details of modeling a power system subject to uncertain power injections, the corresponding generation recourse policy and details of the stochastic OPF formulation.

2.3.2.1 Power System Under Uncertainty

We consider a power network $(\mathcal{B}, \mathcal{L})$. Without loss of generality, for the remaining of this chapter, we assume (i) at most one generator at each location ($|\mathcal{G}_i| = 1,\ i \in \mathcal{B}$), (ii) that the net power injections, D_i^p and D_i^q at each bus $i \in \mathcal{B}$, are subject to uncertainty. Consider an uncertainty realization ω in the possibly unknown/non-parametric set Ω. The set of power flow equations under uncertainty ω is slightly different from (1.5b)-(1.5e) and given by:

$$p_i(\omega) - (D_i^p + \mu_i^p(\omega)) = \sum_{(j=\delta(i)} f_{ij}^p(\omega) + G_i^s v_i^2(\omega), \quad \forall i \in \mathcal{B} \tag{2.11a}$$

$$q_i(\omega) - (D_i^q + \mu_i^q(\omega)) = \sum_{j=\delta(i)} f_{ij}^q(\omega) - B_i^s v_i^2(\omega), \quad \forall i \in \mathcal{B} \tag{2.11b}$$

$$\begin{aligned} f_{ij}^p(\omega) &= -G_{ij} v_i^2(\omega) + G_{ij} v_i(\omega) v_j(\omega) \cos(\theta_i(\omega) - \theta_j(\omega)) \\ &\quad + B_{ij} v_i(\omega) v_j(\omega) \sin(\theta_i(\omega) - \theta_j(\omega)) \quad \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R \end{aligned} \tag{2.11c}$$

$$\begin{aligned} f_{ij}^q(\omega) &= B_{ij} v_i^2(\omega) + G_{ij} v_i(\omega) v_j(\omega) \sin(\theta_i(\omega) - \theta_j(\omega)) \\ &\quad - B_{ij} v_i(\omega) v_j(\omega) \cos(\theta_i(\omega) - \theta_j(\omega)) \quad \forall (i,j) \in \mathcal{L} \cup \mathcal{L}^R \end{aligned} \tag{2.11d}$$

In (2.11), $\mu_i^p(\omega), \mu_i^q(\omega)$ denotes the active and reactive power fluctuations at bus i, under uncertainty ω. All other variables in the system are explicitly expressed as a function of the uncertainty realization.

Recourse Model: For a non-zero realization of uncertainty, the generators in the system must adjust their generation to maintain total power balance

and feasibility. We use an affine policy representing the automatic generation control (AGC) that is representative of current power system operation [ROKA13]. Using affine policies is one of the most straightforward ways to handle uncertainty in power systems and it offers several advantages such as keeping the problem tractable and being easily implementable, among others [ROKA13, BCH14, RMKA16].

$$p_g(\omega) = p_g^0 + \left(\sum_{i\in\mathcal{B}} \mu_i^p(\omega)\right) \alpha_g,\ \forall g \in \mathcal{PV},\ \forall \omega \in \Omega \tag{2.12a}$$

$$v_g(\omega) = v_g^0,\ \forall g \in \mathcal{PV},\ \forall \omega \in \Omega \tag{2.12b}$$

Equation (2.12a) shows the linear adjustment in the active power generation of generator g from its nominal value of p_g^0 as a fraction of the total power mismatch $\sum_{i\in\mathcal{B}} \mu_i^p(\omega)$ caused by the uncertainty, according to its participation factor α_g. In this work, we consider the participation factors to be given and fixed. For simplicity, we assume $\alpha_g = \frac{1}{|\mathcal{G}|}$, although this specific choice is not relevant for our method. Assuming that the participation factor is not a decision variable of the problem is common practice [ROKA13]. Equation (2.12b) says that the voltage magnitudes at PV buses are kept constant during operation, and is in accordance with current practice.

2.3.2.2 Stochastic Optimal Power Flow Formulations

In this section, we present the stochastic optimal power flow problem in a generic form. Since our solution approach involves a Monte Carlo in-the-loop validation step, we have the flexibility to handle a variety of such formulations. We state the set of inequality constraints in the OPF representing the standard safety limits on line flows, phase angle difference at neighboring buses, and bus injections and voltages that need to be enforced.

$$\Gamma_{\text{OPF}} = \{x = (p, q, v, \theta, f^p, f^q) \mid (v, \theta, f^p, f^q) \in \mathcal{OC} \text{ and } (p_g, q_g) \in \mathcal{GC}_g,\ \forall g \in \mathcal{G}\} \tag{2.13a}$$

In the above definition, Γ_{OPF} denotes the set of all power flow solutions that satisfy the safety limits given in (2.13).

Dependent and independent variables: For clarity of exposition, we first specify which variables in the stochastic OPF are controllable/independent and which variables are dependent. Suppose that the nominal values of generation p^0 and voltages v^0 at the PV buses have been determined. Assume that for each realization of the uncertainty ω, the generators react according to the recourse policy in (2.12). Then given ω, Equations (2.12), fully determine the active power generation and voltage magnitude $p_i(\omega), v_i(\omega)$ at all PV buses.

The (known) functions $\mu_i^p(\omega), \mu_i^q(\omega)$ fully determine all real and reactive power injections $p_i(\omega), q_i(\omega)$ at the PQ buses. Once these variables are specified, we are in the standard Power Flow setting, and the set of equations in (2.11) fully specify the value of the rest of the variables – $q_i(\omega), \theta_i(\omega)$ at the PV buses, and $v_i(\omega), \theta_i(\omega)$ at the PQ buses, and all line flows $f_{ij}^p(\omega), f_{ij}^p(\omega)$. We summarize this functional mapping using the following notation:

$$(p(\omega), q(\omega), f(\omega), v(\omega), \theta(\omega)) = \text{PF}\left(p^0, v^0, \omega; \alpha\right). \tag{2.14}$$

A stochastic optimal power flow problem in generic form corresponds to finding a set of nominal set point values for the active power generation p^0 and voltage magnitude v^0 such that a certain cost is minimized, and some stochastic measure of power flow violation for a given uncertainty distribution is below a required limit ϵ. This is made precise in the formulation below:

$$\min_{p^0, v^0} \quad \sum_{g \in \mathcal{G}} C_g(p_g^0) \tag{2.15a}$$

$$\text{s.t.} \quad \mathcal{SV} = \mathbf{E}_{\mathbb{P}_\omega}\left[\mathcal{V}\Big(\text{PF}\left(p_g^0, v^0, \omega; \alpha\right), \Gamma_{\text{OPF}}\Big)\right] \leq \epsilon. \tag{2.15b}$$

Equation (2.15a) specifies the objective that minimizes the total *nominal* generation cost. This is for simplicity. It is possible to incorporate the cost of reserves in a straightforward way. Equation (2.15b) enforces that some *stochastic violation measure* is bounded. The stochastic violation measure $\mathcal{SV}$ is the expectation of some violation measure $\mathcal{V}()$ with respect to $\mathbb{P}_\omega$ which denotes the probability distribution of the uncertainty ω. The violation measure $\mathcal{V}()$ is a function of the uncertainty dependent power flow variables (first argument) and the feasibility/safety region (second argument), and is used to quantify how far the uncertain power flow variables are from the feasible region. Note that the generic formulation in (2.15) includes common cases, such as, *Chance Constrained OPF (CCOPF) [BCH14].* Chance-constrained models in electricity markets are starting to gain attention [CS18] and such constraints are increasingly appearing in practice. A chance-constrained formulation enforces that the probability of constraint violation is smaller than a specified value and corresponds to:

$$\mathcal{V}\Big(\text{PF}\left(p_g^0, v^0, \omega; \alpha\right), \Gamma_{\text{OPF}}\Big) = \mathbb{1}\big(p(\omega), q(\omega), f(\omega), v(\omega), \theta(\omega) \notin \Gamma_{\text{OPF}}\big), \tag{2.16}$$

$$\mathcal{SV} = \mathbb{P}_\omega\Big(p(\omega), q(\omega), f(\omega), v(\omega), \theta(\omega) \notin \Gamma_{\text{OPF}}\Big), \tag{2.17}$$

where $\mathbb{1}()$ denotes the indicator function.

Closed-form analytic expressions for the stochastic constraint in (2.15b) are not easy to derive for the AC-PF model under general uncertainty distributions. To overcome intractability, data driven scenario OPF can be formulated.

2.3.2.3 Scenario OPF (S-OPF)

Scenario approach [CC06, VMLA13] collects a set Ω_N of N random samples for the uncertainty $\omega \in \Omega$. By definition, the base case $\omega = 0$ is included in set Ω_N, and the user is assumed to have access to a scenario generation/sampling process (from historical data or otherwise). We then solve an OPF problem with hard feasibility constraints for each selected scenario as denoted below.

$$\min_{p^0, v^0} \quad \sum_{g \in \mathcal{G}} C_g(p_g^0) \tag{2.18a}$$

$$\text{s.t.} \quad \forall \omega_i \in \Omega_N, \ \text{PF}\left(p_g^0, v^0, \omega_i; \alpha\right) \in \Gamma_{\text{OPF}} \tag{2.18b}$$

By ensuring feasibility for a large-enough and representative sample set Ω_N, S-OPF can indirectly guarantee the stochastic violation constraint (2.15b). Problem (2.18) provides a robust formulation if one were to consider Ω_N as the uncertainty set. A careful choice of the scenarios to be added to Ω_N leads to a solution of the stochastic version (2.15) on the original set Ω. This is what is experimentally demonstrated in this chapter. Theoretical bounds on the size of the sample set necessary to ensure $\mathcal{SV}() \leq \epsilon$ and related design of box constraints exist for convex optimization problems [CC06, MGL14], but are not generalizable to AC-OPF. As demonstrated later, the number of samples to ensure low stochastic violation quickly grows. This makes the standard S-OPF in (2.18) computationally intractable for realistically sized test cases. Existing scenario selection methods pick a sub-set of scenarios from the ones available, randomly [CDT$^+$15] or by minimizing an inter-distribution distance such as Wasserstein metric [GKHR03]. Similarly, mixed-integer programs have been proposed to pick a sub-set of scenarios inside chance-constrained optimization [DGKR03]. However the number of selected scenarios necessary, or the mixed-integer programs themselves, still involve a large computational requirement for AC- OPF. In this work, we take a different approach where system knowledge and data-driven techniques are combined to design (not just select) strategic scenarios that lead to a drastically more efficient scenario OPF.

2.3.3 Data-Driven Scenario OPF

The overarching goal of our approach is to determine an optimized scenario set Ω_N of far lesser cardinality, compared to random sampling, so that a tractable scenario OPF solution with stochastic violations below prescribed threshold can be be determined. We propose an algorithm called *Data-Driven Scenario OPF* (DDS-OPF):

Algorithm 2 DDS-OPF

1: **Initialization:** Solve S-OPF (2.18) using rated loads D_i^p, D_i^q ($|\Omega_N| = 1$) to get (p^0, v^0);
2: **Monte Carlo:** Sample a set of permissible scenarios $\mathcal{S}$ of size S according to $\mathbb{P}_\omega$. Solve the PF with recourse for each scenario;
3: **Stopping criterion check:** Check if the *estimated* stochastic violation measure $\tilde{\mathcal{SV}}$ is below pre-selected threshold, $\tilde{\mathcal{SV}} = \frac{1}{S}\sum_{i=1}^{S} \mathcal{V}(*, \omega_i) < \tau$. If yes, **exit**;
4: **Scenario construction:** Use data-driven methods to design $K < S$ scenarios to add to Ω_N;
5: **Update:** Compute new solution (p^0, v^0) for (2.18) with Ω_N. Go to Step **Monte Carlo**;
6: **return** p^0, v^0.

Note that we sample a new set $\mathcal{S}$ each time the **Monte Carlo** step is executed. The threshold τ used in DDS-OPF is selected based on the properties of the stochastic violation measure $\mathcal{SV}()$, pre-fixed ϵ bound (see (2.15b)), and the confidence requirement. Theoretical confidence bound on the solution for selected τ is given in Section 2.3.3.3. All numerical experiments focus on the case when $\mathcal{SV}()$ corresponds to the probability of constraint violation. For that, the estimated stochastic violation measure $\tilde{\mathcal{SV}}$ simply corresponds to the fraction of samples in $\mathcal{S}$, for which the constraints are violated. Note that the success of Algorithm 2 depends on the solvability of (2.18) in the **Update** step. Indeed, if Ω_N is too large, problem (2.18) becomes untractable. This is why step **Scenario construction** is crucial to keep Ω_N at a reasonable size and this is the main focus of our discussion in the next paragraphs.

The rest of the section is focused on describing the critical Step **4** in DDS-OPF. To guide intuition, we use computations on the `pglib_opf_case73_ieee_rts` test-case in the OPF Power Grid Library [BBC+19]. This case has 73 buses, 120 lines and 51 loads. We assume that $\mathbb{P}_\omega$ is a uniform distribution within a box, i.e., for each load $i \in \mathcal{B}$ we have $\mu_i^p(\omega) \sim \mathcal{U}[-0.03D_i^p, 0.03D_i^p]$ and $\mu_i^q(\omega) \sim \mathcal{U}[-0.03D_i^q, 0.03D_i^q]$, where $\mathcal{U}$ denotes the uniform distribution, and D_i^p, D_i^q are the rated active and reactive demands.

Table 2.4 shows the performance of the vanilla scenario approach where randomly drawn samples are included in the set Ω_N[5]. Note that a non-trivial number of violations are still obtained despite 50 scenarios. This high sample requirement prevents tractability for realistic test-cases. Note also that since we rely on a box uniform distribution, we can compute the solution with 5 scenarios which are the central scenario and the four extreme scenarios (each

[5]The results are obtained for only one run of samples and explains how on this run having 20 scenarios performs better than 30 (same with 40 and 50).

Table 2.4: Feasibility on 1,000 out-of-sample scenarios for DDS-OPF with randomly sampled Ω_N with $K = 10$, for 73-bus test system.

$\lvert\Omega_N\rvert - 1$	1	10	20	30	40	50
P_{vio}^{1000}	100%	59.5%	25.0%	32.3%	8.00%	12.2%
Cost	1.904e5	1.948e5	1.948e5	1.948e5	1.948e5	1.948e5

corner of the box). For this scenario set, the infeasibility rate reaches 16.0% which is still high.

To improve over random sampling, our proposed *scenario construction* in Step **4** includes 2 key sub-steps:

(a) *PF-aware scenario selection*: We use prioritization metrics to down-select *dominant* scenarios.

(b) *Data-driven scenario enhancement*: For scenarios selected in (a), we identify critical directions that maximize their effect on S-OPF, and modify them (stretch or squeeze) along these directions before adding to Ω_N.

A schematic representation of our overall approach is shown in Fig. 2.2. In what follows, we describe in detail, the motivation and important features of scenario construction sub-steps and use the 73-bus test system to demonstrate improvements.

2.3.3.1 PF-Aware Scenario Selection

A random scenario, that is already feasible for the current solution (p^0, v^0), is less likely to be effective for feasibility improvement than a scenario that has multiple constraint violations during recourse. We use information about constraint violations to sub-select a small number of *dominant* scenarios from set $\mathcal{S}$ in Step **2** to add to the scenario set Ω_N. Fortunately, the infeasible scenarios and their corresponding constraint violations are already acquired while validating the performance of (p^0, v^0) in Step **3**.

Dominant scenario selection. We consider three different prioritization criteria:

- *Maximum violation (MV).* Scenarios having the largest constraint violation, measured relative to bound value.

- *Number of constraints (NC).* Scenarios violating the maximum number of constraints.

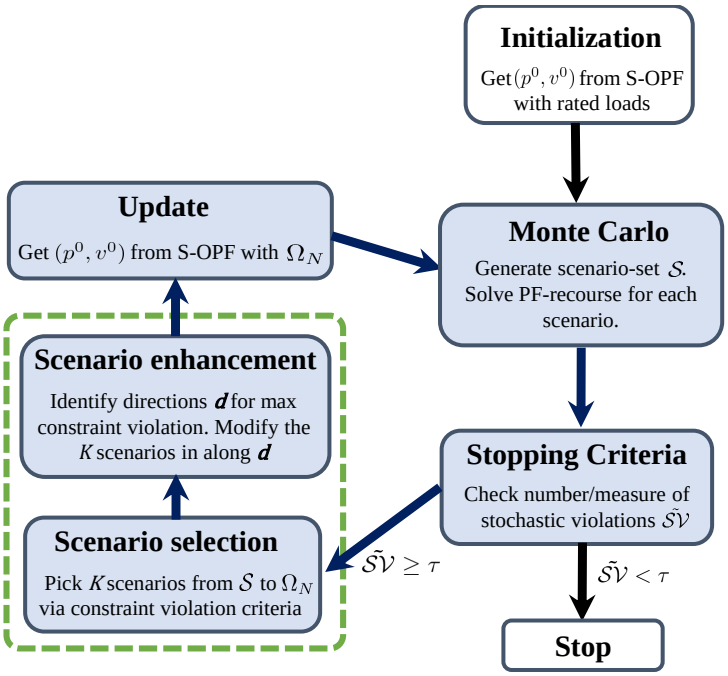


Figure 2.2: Schematic of DDS-OPF. The scenario construction sub-steps are highlighted within the green box.

- *Hybrid.* Scenarios that have the highest $weight_s = \frac{MV_s}{\max\limits_{s' \in \mathcal{S}} MV_{s'}} + \frac{NC_s}{\max\limits_{s' \in \mathcal{S}} NC_{s'}}$, where MV_s is the largest violation of a constraint, and NC_s is the number of constraints violated, by scenario s.

We avoid selecting a new scenario that violates the same set of constraints as a previously selected (dominating) scenario. Such avoidance ensures that a greater proportion of constraint violations are represented in Ω_N.

Batch size selection. While the prioritization criteria rank the scenarios according to their dominance, the number of samples K that are added back to Ω_N still needs to be decided and can have a significant impact on overall efficiency. When K is too small the total number of iterations can be large since we are adding very little information to the problem in each iteration. On the other hand, when K is too large, the size of the resulting S-OPF can quickly make it intractable and there is the risk of 'oversatisfying' the probabilistic constraint (2.15b). These observations are confirmed in Figure 2.3 for the 73-bus system. Based on these observations, we decide that a batch size of $K = 5$ provides the right trade-off across a variety of test cases (later confirmed

in section 2.3.4). Note that the number of final scenarios is often less than $\#iterations * K + 1$, since in each iteration, only one of multiple scenarios that violate the same set of constraints, is added to Ω_N. In other words, if a given set of constraints is violated by multiple samples, we only need to add one of them to Ω_N. This feature is analogous to observations in [NMRB18, DM19] on a sparse set of active constraints in OPF.

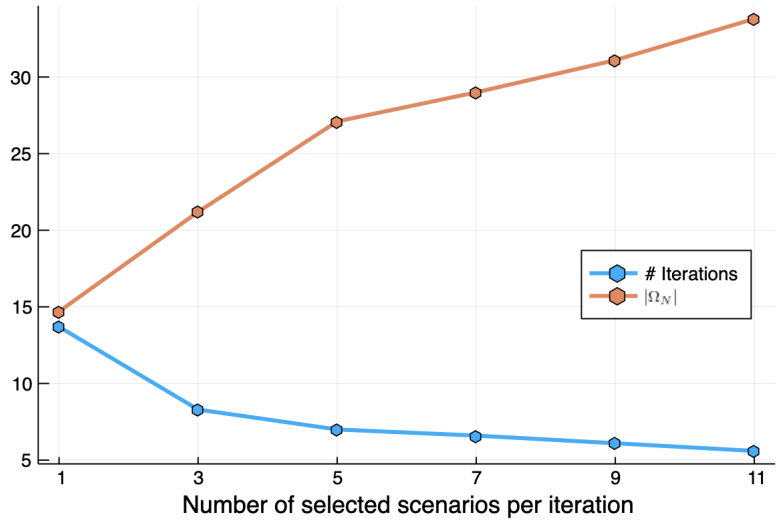


Figure 2.3: Number of iterations and final size of Ω_N when the algorithm has converged for different choices for K, number of selected scenarios per iteration, for the 73-bus system. These are average numbers out of 10 runs of DDS-OPF.

Results for 73-bus case. We use dominant scenario selection in DDS-OPF with $S = 1000$, $K = 5$ and $\tau = 0$ and show the results in Table 2.5. Compared to the results for the random sampling in Table 2.4, we have signif-

Table 2.5: Feasibility on $1,000$ out-of-sample scenarios for DDS-OPF with scenario selection with $K = 5$, for 73-bus test system.

Policy	# Iterations	$\lvert\Omega_N\rvert$	P_{vio}^{1000}
MV	5	20	0.1 %
NC	7	28	0 %
Hybrid	8	29	0 %

icantly improved performance for each of the three proposed criteria. Indeed,

at most 29 scenarios are able to reduce the number of infeasible scenarios to almost zero on out-of-sample testing[6].

2.3.3.2 Data-Driven Scenario Enhancement

Note that while Section 2.3.3.1 allows us to select scenarios through prioritization metrics, we do not modify the generated scenarios. In this section, we present data-driven enhancements to selected scenarios before adding them to Ω_N, that make our approach more efficient and amenable for large test-cases. Based on preliminary tests on multiple cases, we observe that violations of a given constraint are primarily caused by a small subset of load fluctuations. Further, there are certain *critical directions* for these load fluctuations that maximize violation. We now describe our method to identify this subset of loads and the critical directions, and a procedure to enhance the selected scenarios along these critical directions to make them more effective in enforcing feasibility.

Identifying critical components and directions. Our approach to critical component identification relies on regularized linear regression [WJ+08], as described next. Consider selected scenario $t = (\mu^p(t), \mu^q(t))$ that we intend to enhance. Let $\mathcal{C}_t$ be the set of constraints violated by t during recourse. For each $c \in \mathcal{C}_t$, let $\mathcal{S}_c$ ($\subseteq \mathcal{S}$) be the set of random samples that violate constraint c, and define the relative violation u_c^s for sample $s = (\mu^p(s), \mu^q(s))$. We approximate a sparse linear map between the active and reactive load fluctuations in buses $\mathcal{B}$, and violation for constraint $c \in \mathcal{C}_t$. The critical components and directions are identified via the vector $d_c = \{d_0\} \cup \{d_i, i \in \mathcal{B}\}$, computed as follows:

$$d_c = \arg\min_d \sum_{s \in \mathcal{S}_c} \left(u_c^s - (d_0 + \sum_{\substack{i \in \mathcal{B} \\ r=(p,q)}} d_i^r \mu_i^r(s)) \right)^2 + \lambda |d|_1.$$

Here $\lambda > 0$ is a regularization coefficient used with the ℓ_1 norm to promote sparse solutions. This is an unconstrained convex optimization problem that can be easily solved, including in parallel for each selected scenario t and constraint c.

Scenario enhancement. Using the critical directions identified, we describe the scenario enhancement procedure for the special case when the uncertainty is a uniform distribution over a box. There are variations possible

[6] We do not increase the number samples because we explain the main steps of our methodology on the 73-bus system before testing larger test systems.

for other distributions, which we do not pursue in the work. The enhancement operation for scenario t is given below:

$$\forall i \in \mathcal{B},\ r = (p, q), \qquad \text{if } |d_i^r| < \tau_2 \text{ then } \mu_i^r(t) \leftarrow \mu_i^r(t)$$
$$\text{else } \mu_i^r(t) \leftarrow \begin{cases} \overline{\mu_i^r} & \text{if } d_i^r > \tau_2. \\ \underline{\mu_i^r} & \text{if } d_i^r < -\tau_2. \end{cases}$$

where $\tau_2 > 0$ is a positive threshold.

Note that the enhancement step changes entries in scenario t to their maximum or minimum values, based on the sign of non-trivial entries in d_c. This is done as the signs in d_c reflect positive or negative directions to maximize violation. In settings where the maximum values of μ_i^p, μ_i^q are not known, one can change it by a factor of the current entries (akin to a gradient based change). In this work, we use $\tau_2 = 1e^{-4}$ for our simulations. By increasing the threshold τ_2, the changes in t can be made more sparse.

Results for 73-bus case. In addition to scenario selection of Section 2.3.3.1, we now use the scenario enhancement technique on the 73-bus test case. The results are presented in Table 2.6. We observe that the addition of scenario enhancement significantly reduces (more than 60%) the number of samples necessary for convergence of DDS-OPF. The combined impact of scenario selection and scenario enhancement steps over random sampling is evident from comparisons with Table 2.4. Using at most 11 optimized scenarios, our proposed method is able to bring down infeasibility in out-of sample testing from 50% to 0.

Table 2.6: Feasibility on $1,000$ out-of-sample scenarios for DDS-OPF with scenario selection & scenario enhancement with $K = 5$, for 73-bus test system.

Policy	# Iterations	$\lvert\Omega_N\rvert$	P_{vio}^{1000}
MV	1	6	0 %
NC	2	11	0 %
Hybrid	2	11	0 %

2.3.3.3 Monte-Carlo Step, Confidence Bounds, and Scaling

In this section, we provide a theoretical confidence bound on the quality of the solution obtained from DDS-OPF based on the stopping criterion τ employed in step **3**. The proof relies on an application of the Hoeffding inequality [WJ+08].

Theorem 2.3.1. *Suppose that for all nominal power flow solutions in Γ_{OPF} and for all $\omega \in \Omega$, the violation measure satisfies $|\mathcal{V}()| \leq M$. Then the solution*

(p^0, v^0) *obtained from* DDS-OPF *with stopping criterion* τ *and sample size* S *satisfies*

$$\mathbb{P}_\omega\left(\mathcal{SV} < \tau + \alpha S^{-1/2}\right) > 1 - \delta, \quad \textit{where} \;\; \alpha = \sqrt{2M^2 \log(1/\delta)}.$$

Proof. Since $\mathcal{V}(*, \omega)$ is a random variable as a function of uncertainty realization ω bounded by M (the dependence on other non-random quantities has been suppressed for clarity). By using the Hoeffding inequality [WJ+08] for (2.15b), we get for any $t > 0$,

$$\mathbb{P}_\omega\left(\mathcal{SV} > \frac{1}{S}\sum_{i=1}^{S} \mathcal{V}(*, \omega_i) + t\right) \leq \exp\left(-St^2/2M^2\right).$$

The proof follows by using $t = \alpha S^{-1/2}$. □

Theorem 2.3.1 shows how the stopping criterion translates to the quality of the solution. A critical advantage of DDS-OPF is the Monte-Carlo-in-the-loop step **2**. This is different from the vanilla scenario approach, where the random samples drawn from $\mathbb{P}_\omega$ are incorporated into S-OPF. In contrast, in DDS-OPF the samples used in step **2** to evaluate the current solution p^0, v^0 are *independent* from the samples used in the prior iteration to obtain p^0, v^0 (step **5** or step 1 in the first iteration). This results in fast convergence rates obtained via Theorem 2.3.1.

In all our experiments in Section 2.3.4, we choose $S = 1000$ and $\tau = 0$ with $\mathcal{SV}$ = probability of violation. Since the probability is always smaller than 1, we have $M = 1$. By applying Theorem 2.3.1, we can guarantee with confidence 95%, that all solutions obtained in this work satisfy the joint chance constraints with probability 99%. This reliability level is in line with standard reliability targets envisioned in EU legislation and among European TSOs in practice [otNsa18, DVSD+19].

The Monte-Carlo step involves solving a series of power flows. Since the loading conditions resulting from uncertainty are still in the vicinity of the nominal load, warm-start methods can be used to solve a large number of power flows quickly. Further, this easily lends itself to parallelization, resulting in even further reduction in computation time. As a result, most of the computational complexity of DDS-OPF lies in solving the resulting S-OPF in Step **5**.

2.3.4 Numerical Experiments

In this section, we benchmark the DDS-OPF by detailed numerical experiments on a number of test cases in the IEEE PES PGLib-OPF benchmark library. The code is accessible from the following link : `https://github.com/imezghani/StochasticACOPF`.

2.3.4.1 Test Cases and Experiment Setup

We consider four different test cases, `24_ieee`, `73_ieee`, `118_ieee` and `1354_pegase`. The details of the test cases are shown in Table 2.7. For the first three (smaller) test cases, we assume that all active and reactive loads have a uniform 3% fluctuation around their nominal value. For the `1354_pegase` test case, we assume that the real and reactive powers of the 211 out of the 673 loads that are situated at end-buses fluctuate uniformly by 2% of their nominal value. These buses often correspond to connections to distribution/sub-transmission, where the consumers and distributed energy resources responsible for the uncertainty are situated. The network is illustrated in Fig. 2.4. From Table 2.7, it is clear that the recourse with the base-case solution can lead to infeasibility for an extremely high number of load fluctuations ($> 85\%$).

We remark here that the level of uncertainty chosen is quite large; increasing the uncertainty further from the given values makes a large percentage of loading conditions infeasible for the basic OPF, let alone the stochastic OPF. For DDS-OPF, we choose $S = 1000$, $K = 5$ and $\tau = 0$ with empirical probability of violation $\tilde{\mathcal{SV}}$.

Table 2.7: Test case details

Test case	`24_ieee`	`73_ieee`	`118_ieee`	`1354_pegase`
# Buses	24	73	118	1,354
# Generators	33	99	54	260
# Lines	38	120	186	1,991
# Loads	17	51	99	673
# Fluctuations	17	51	99	211
Base cost	$6.34e4$	$1.90e5$	$9.72e4$	$1.26e6$
Base P_{vio}^{1000}	87.5 %	100%	100%	100%

2.3.4.2 Performance Trends

Table 2.8 shows the results of applying DDS-OPF on the different test cases.

Scenario size. We see that DDS-OPF has excellent performance for all test cases in terms of number of iterations ($\#It$) and final number of samples ($|\Omega_N|$). $|\Omega_N|$ grows very slowly with network size, from 7 on the 24 bus system to only 31 on the 1354 bus system. This demonstrates that DDS-OPF has very favorable scaling properties, and can be scaled to even larger systems.

Cost. While the scenario enhancement procedure introduced in Section 2.3.3.2 helps quickly obtain a secure solution, worsening the scenarios can potentially

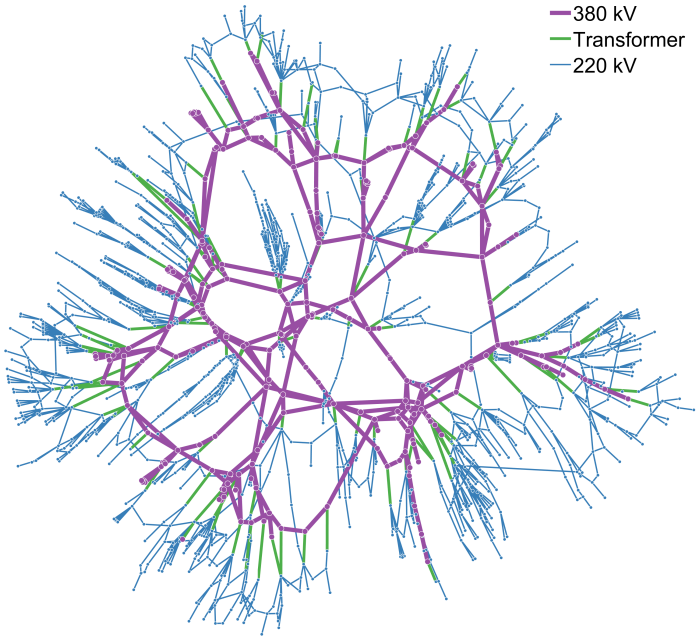


Figure 2.4: Topology of test case `1354_pegase`.

increase the cost. However, we see by comparing the cost between Table 2.7 and Table 2.8 that in all the cases the increase in cost from the deterministic (and unsafe) solution is small with $\sim 2\%$ for the 24 bus system to $\sim 0.2\%$ for the 1354 bus system. The larger, more realistic system, possesses more flexibility to handle uncertainty in a more economic way, as expected.
Furthermore, by comparing the costs for the 73-bus system in Tables 2.4 and 2.8, it is clear that our algorithm achieves the same cost as the vanilla scenario selection scheme, while significantly improving the feasibility of the solution.

Distance to the deterministic solution. The last two columns of Table 2.8 report the $2-$norm difference between the deterministic solution set-points and the DDS-OPF solution set-points, first in terms of real power injections[7] and then voltage magnitudes. The quantities suggest that the solution to the stochastic OPF lies in the vicinity of the solution of the deterministic OPF. Nevertheless, this adjustment to the deterministic solution is critical and can significantly improve the robustness of the solutions. Using the 1354 bus system as an example, the reduction in maximum violation can be as much as 17% (see Table 2.9).

[7]The measure is in per unit: 1 p.u. = 100 MW

Table 2.8: Overall performance trends of DDS-OPF

Test case	Policy	# It	$\|\Omega_N\|$	P_{vio}^{1000}	Cost	Dist P	Dist V
`24_ieee`	MV	3	7	0 %	6.502e4	$2.5e{-}1$	$4.4e{-}3$
`73_ieee`	MV	1	6	0 %	1.948e5	$5.4e{-}1$	$9.9e{-}3$
`118_ieee`	Hybrid	3	14	0 %	9.802e4	$9.0e{-}1$	$3.7e{-}2$
`1354_pegase`	MV	6	31	0.1 %	1.263e6	1.4	$5.1e{-}2$

2.3.4.3 A Detailed Study on the `1354_pegase` Test Case

We present detailed numerical experiments for different variants of DDS-OPF on the 1354 bus system. Table 2.9 shows the results for different choices of prioritization rule.

Effect of prioritization rule. With any prioritization rule, DDS-OPF finds an excellent solution with a maximum of 31 scenarios in the final S-OPF, NC being slightly less performant on this particular test case. All resulting costs are similar, and within 0.2% of the base case cost.

Different stochastic violation measures. The Monte Carlo in-the-loop method employed by DDS-OPF grants it the flexibility to handle a variety of stochastic violation measures. Table 2.9 shows two such violation measure, the *probability of violation* and *maximum magnitude of violation* in an out of sample testing with 1000 samples. This translates into confidence guarantees in the sense of Theorem 2.3.1. As an example by using Theorem 2.3.1, we can guarantee that the solution obtained using MV for constraint selection, satisfies a chance constraint with probability of violation $< 1.1\%$. Similarly, we can guarantee that in the face of uncertainty, the solution has a maximum constraint violation of 3.26%. The second guarantee uses a very conservative maximum violation bound of $M = 10$. Both the above statements carry a confidence of 95%.

Table 2.9: Results of the iterative approach on `1354_pegase`.

Policy	# It	$\|\Omega_N\|$	P_{vio}^{1000}	Max. Viol.	Cost ($\times 1e6$)
Base	-	1	100 %	17.3 %	1.2620
MV	6	31	0.1 %	0.06 %	1.2633
NC	6	31	2.3 %	0.34 %	1.2633
Hybrid	8	31	0.1 %	0.04 %	1.2634

2.3.5 Conclusion and Future Directions

This work describes a principled iterative data-driven approach for stochastic AC-OPF under general probabilistic constraints. The non-linear and non-convex equations in AC-OPF make random sampling or scenario reduction approaches impractical for large test cases, due to their large sample requirement. Our data-driven algorithm is able to overcome that by a novel 2-step process for 'dominant' scenario design/construction that involves: (a) scenario selection based on constraint violations, and (b) scenario enhancement by regularized linear regression. Through system-level intuition, theoretical bounds, and finally numerical verification on multiple test cases, we demonstrate that our data-driven algorithm is able to provide feasible solutions to stochastic AC-OPF using far lower scenarios than conventional schemes. For example, our method uses only 31 constructed samples to provide a feasible solution for the `1354_pegase` test case, that satisfies chance constraints with $< 1.1\%$ violation probability.

This work naturally leads to multiple extensions. First, we plan to parallelize the steps (scenario enhancement, Monte Carlo checks) and include warm-starts in our algorithm to reduce and distribute the computational effort. While the current work operates on box-uncertainty sets for sampling and scenario enhancement, efficient data-driven efforts for general (non-parametric) uncertainty sets is another direction for exploration. Finally we plan to analyze extensions of our approach to related and computationally challenging problems on resilient network design and stochastic unit commitment. For example, the case of $(N-1)$ security constrained optimal power flow could be of interest. Our framework could easily be adapted if one were to consider lines as contingencies[8] where the uncertainty would be due to the possible contingency of a line instead of the realized demand.

[8]The case of possible failures of generators might be more difficult to handle since we use an affine policy to make all generators react to possible disturbances.

Appendix

2.A Gauss-Newton Algorithm Convergence Analysis

2.A.1 Global Convergence Analysis

We first summarize some properties of Algorithm 1 when the penalty parameter β is fixed at a given positive value for all iterations.

Lemma 2.A.1. *Let $\boldsymbol{V}_L$ be defined by (2.6), and $\boldsymbol{G}_L$, $\boldsymbol{d}_L$, and r_L be defined by (2.7). Then the following statements hold:*

(a) *If $\boldsymbol{V}_{L_k}(\boldsymbol{x}^k) = \boldsymbol{x}^k$, then $\boldsymbol{x}^k$ is a stationary point of (2.2).*

(b) *The norm $\|\boldsymbol{G}_{L_k}(\boldsymbol{x}^k)\|$ is nondecreasing in L_k, and $r_{L_k}(\boldsymbol{x}^k)$ is nonincreasing in L_k. Moreover, we have*

$$F(\boldsymbol{x}^k) - \mathcal{Q}_{L_k}(\boldsymbol{V}_{L_k}(\boldsymbol{x}^k); \boldsymbol{x}^k) \geq \frac{L_k}{2} r_{L_k}^2(\boldsymbol{x}^k). \tag{2.19}$$

(c) *If $\Psi'(\cdot)$ is Lipschitz continuous with the Lipschitz constant L_Ψ, then, for any $x \in \mathcal{X}$, we have*

$$\begin{aligned} &F(\boldsymbol{x}^k) - F(\boldsymbol{V}_{L_k}(\boldsymbol{x}^k)) \geq \tfrac{(2L_k - \beta L_\Psi)}{2} r_{L_k}^2(\boldsymbol{x}^k) = \tfrac{(2L_k - \beta L_\Psi)}{2L_k^2} \|\boldsymbol{G}_{L_k}(\boldsymbol{x}^k)\|^2. \\ &DF(\boldsymbol{x}^k)[\boldsymbol{d}_{L_k}(\boldsymbol{x}^k)] \leq -L_k r_{L_k}^2(\boldsymbol{x}^k) = -\tfrac{1}{L_k} \|\boldsymbol{G}_{L_k}(\boldsymbol{x}^k)\|^2. \end{aligned} \tag{2.20}$$

Proof. (a) Substituting $\boldsymbol{V}_{L_k}(\boldsymbol{x}^k) = \boldsymbol{x}^k$ into (2.8), we again obtain the optimality condition (2.4). This shows that $\boldsymbol{x}^k$ is a stationary point of (2.2).

(b) Since the function $q(t, \boldsymbol{x}) := f(\boldsymbol{x}) + \beta|\Psi(\boldsymbol{x}^k) + \Psi'(\boldsymbol{x}^k)(\boldsymbol{x} - \boldsymbol{x}^k)| + \frac{1}{2t}\|\boldsymbol{x} - \boldsymbol{x}^k\|^2$ is convex in two variables $\boldsymbol{x}$ and t, we have that $\eta(t) := \min_{\boldsymbol{x} \in \mathcal{X}} q(t, \boldsymbol{x})$ is still convex. It is easy to show that $\eta'(t) = -\frac{1}{2t^2}\|\boldsymbol{V}_{1/t}(\boldsymbol{x}^k) - \boldsymbol{x}^k\|^2 = -\frac{1}{2t^2}\|\boldsymbol{d}_{1/t}(\boldsymbol{x}^k)\|^2 = \frac{1}{2}\|\boldsymbol{G}_{1/t}(\boldsymbol{x})\|^2$. Since $\eta(t)$ is convex, $\eta'(t)$ is nondecreasing in t. This implies that $\|\boldsymbol{G}_{1/t}(\boldsymbol{x}^k)\|$ is nonincreasing in t. Thus $\|\boldsymbol{G}_L(\boldsymbol{x}^k)\|$ is nondecreasing in L and $r_L(\boldsymbol{x}^k) := \|\boldsymbol{d}_L(\boldsymbol{x}^k)\|$ is nonincreasing in L. To prove (2.19), note that the convexity of η implies that

$$F(\boldsymbol{x}^k) = \eta(0) \geq \eta(t) + \eta'(t)(0 - t) = \eta(t) + \frac{1}{2t} \boldsymbol{r}_{1/t}^2(\boldsymbol{x}^k). \tag{2.21}$$

On the other hand, $\mathcal{Q}_L(\boldsymbol{V}_L(\boldsymbol{x}^k); \boldsymbol{x}^k) = \eta(1/L)$. Substituting this relation into (2.21), we obtain (2.19).

(c) Let use define $\boldsymbol{V}_k := \boldsymbol{V}_{L_k}(\boldsymbol{x}^k)$. From the optimality condition (2.8), for any $\boldsymbol{x} \in \mathcal{X}$, we have

$$\left[\nabla f(\boldsymbol{V}_k) + L_k(\boldsymbol{V}_k - \boldsymbol{x}^k) + \beta\Psi'(\boldsymbol{x}^k)\xi(\boldsymbol{x}^k)\right]^\top (\boldsymbol{x} - \boldsymbol{V}_k) \geq 0,$$

where $\xi(\boldsymbol{x}^k) \in \partial|\Psi(\boldsymbol{x}^k) + \Psi'(\boldsymbol{x}^k)(\boldsymbol{V}_k - \boldsymbol{x}^k)|$. Substituting $\boldsymbol{x} = \boldsymbol{x}^k$ into this condition, we have

$$\nabla f(\boldsymbol{V}_k)^\top(\boldsymbol{x}^k - \boldsymbol{V}_k) + \beta\xi(\boldsymbol{x}^k)^\top\Psi'(\boldsymbol{x}^k)^\top(\boldsymbol{x}^k - \boldsymbol{V}_k) \geq L_k\|\boldsymbol{V}_k - \boldsymbol{x}^k\|^2. \quad (2.22)$$

Since f is convex, we have:

$$f(\boldsymbol{x}^k) \quad \geq f(\boldsymbol{V}_k) + \nabla f(\boldsymbol{V}_k)^\top(\boldsymbol{x}^k - \boldsymbol{V}_k).$$

By exploiting the convexity of $|\cdot|$ at point $\Psi(\boldsymbol{x}^k) + \Psi'(\boldsymbol{x}^k)(\boldsymbol{V}_k - \boldsymbol{x}^k)$, we have:

$$|\Psi(\boldsymbol{x}^k)| \quad \geq |\Psi(\boldsymbol{x}^k) + \Psi'(\boldsymbol{x}^k)(\boldsymbol{V}_k - \boldsymbol{x}^k)| + \xi(\boldsymbol{x}^k)^\top\Psi'(\boldsymbol{x}^k)^\top(\boldsymbol{x}^k - \boldsymbol{V}_k).$$

Since Ψ' is Lipschitz continuous, we also have

$$\begin{aligned} |\Psi(\boldsymbol{V}_k)| \quad &\leq |\Psi(\boldsymbol{x}^k) + \Psi'(\boldsymbol{x}^k)(\boldsymbol{V}_k - \boldsymbol{x}^k)| + |\Psi(\boldsymbol{V}_k) - \Psi(\boldsymbol{x}^k) + \Psi'(\boldsymbol{x}^k)(\boldsymbol{V}_k - \boldsymbol{x}^k)| \\ &\leq |\Psi(\boldsymbol{x}^k) + \Psi'(\boldsymbol{x}^k)(\boldsymbol{V}_k - \boldsymbol{x}^k)| + \tfrac{L_\Psi}{2}\|\boldsymbol{V}_k - \boldsymbol{x}^k\|^2. \end{aligned}$$

Combining these three bounds, we can show that

$$\begin{aligned} f(\boldsymbol{x}^k) + \beta|\Psi(\boldsymbol{x}^k)| \quad &\geq f(\boldsymbol{V}_k) + \beta|\Psi(\boldsymbol{x}^k) + \Psi'(\boldsymbol{x}^k)(\boldsymbol{V}_k - \boldsymbol{x}^k)| + L_k\|\boldsymbol{V}_k - \boldsymbol{x}^k\|^2 \\ &\geq f(\boldsymbol{V}_k) + \beta|\Psi(\boldsymbol{V}_k)| + L_k\|\boldsymbol{V}_k - \boldsymbol{x}^k\|^2 - \tfrac{\beta L_\Psi}{2}\|\boldsymbol{V}_k - \boldsymbol{x}^k\|^2, \end{aligned}$$

which implies

$$F(\boldsymbol{x}^k) \quad \geq F(\boldsymbol{V}_k) + \tfrac{(2L_k - \beta L_\Psi)}{2}\|\boldsymbol{V}_k - \boldsymbol{x}^k\|^2.$$

Since $r_{L_k}^2(\boldsymbol{x}^k) = \|\boldsymbol{V}_k - \boldsymbol{x}^k\|^2 = \frac{1}{L_k^2}\|\boldsymbol{G}_{L_k}(\boldsymbol{x}^k)\|^2$, we obtain the first inequality of (2.20) from the last inequality. Moreover, from (2.3) we have

$$DF(\boldsymbol{x}^k)[\boldsymbol{d}_{L_k}(\boldsymbol{x}^k)] = \nabla f(\boldsymbol{V}_k)^\top(\boldsymbol{V}_k - \boldsymbol{x}^k) + \beta\xi(\boldsymbol{x}^k)^\top\Psi'(\boldsymbol{x}^k)^\top(\boldsymbol{V}_k - \boldsymbol{x}^k).$$

Using (2.22), we can show that $DF(\boldsymbol{x}^k)[\boldsymbol{d}_{L_k}(\boldsymbol{x}^k)] \leq -L_k\|\boldsymbol{V}_k - \boldsymbol{x}^k\|^2$, which is the second inequality of (2.20). □

The proof of Lemma 2.A.1(a) shows that if we can find $\boldsymbol{x}^k$ such that $\|\boldsymbol{G}_{L_k}(\boldsymbol{x}^k)\| \leq \varepsilon$, then $\boldsymbol{x}^k$ is an approximate stationary point of (2.2) within the accuracy ε. From statement (b), we can see that if the line-search condition $F(\boldsymbol{V}_{L_k}(\boldsymbol{x}^k)) \leq \mathcal{Q}_{L_k}(\boldsymbol{V}_{L_k}(\boldsymbol{x}^k); \boldsymbol{x}^k)$ at Step 4 holds, then $F(\boldsymbol{V}_{L_k}(\boldsymbol{x}^k)) \leq F(\boldsymbol{x}^k) - \frac{L_k}{2}r_{L_k}^2(\boldsymbol{x}^k)$. That is, the objective value $F(\boldsymbol{x}^k)$ decreases at least by $\frac{L_k}{2}r_{L_k}^2(\boldsymbol{x}^k)$ after the k-th iteration. We first claim that Algorithm 1 is well-defined.

Lemma 2.A.2. *Algorithm 1 is well-defined, i.e. step 4 terminates after a finite number of iterations. That is, if $L \geq \beta L_\Psi$, then $F(\boldsymbol{V}_L(\boldsymbol{x}^k)) \leq \mathcal{Q}_L(\boldsymbol{V}_L(\boldsymbol{x}^k); \boldsymbol{x}^k)$.*

Proof. Since Ψ' is L_Ψ-Lipschitz continuous, for any $\boldsymbol{x}^k$ and $\boldsymbol{V}_L(\boldsymbol{x}^k)$, we have

$$\begin{aligned}|\Psi(\boldsymbol{V}_L(\boldsymbol{x}^k))| &\leq |\Psi(\boldsymbol{x}^k) + \Psi'(\boldsymbol{x}^k)(\boldsymbol{V}_L(\boldsymbol{x}^k) - \boldsymbol{x}^k)| \\ &\quad + \|\Psi(\boldsymbol{V}_L(\boldsymbol{x}^k)) - \Psi(\boldsymbol{x}^k) - \Psi'(\boldsymbol{x}^k)(\boldsymbol{V}_L(\boldsymbol{x}^k) - \boldsymbol{x}^k)\| \\ &\leq |\Psi(\boldsymbol{x}^k) + \Psi'(\boldsymbol{x}^k)(\boldsymbol{V}_L(\boldsymbol{x}^k) - \boldsymbol{x}^k)| + \tfrac{L_\Psi}{2}\|\boldsymbol{V}_L(\boldsymbol{x}^k) - \boldsymbol{x}^k\|^2.\end{aligned}$$

Using the definition of $\mathcal{Q}_L(\boldsymbol{V}; \boldsymbol{x})$, we obtain

$$F(\boldsymbol{V}_L(\boldsymbol{x}^k)) \leq \mathcal{Q}_L(\boldsymbol{V}_L(\boldsymbol{x}^k) l \boldsymbol{x}^k) - \tfrac{L - \beta L_\Psi}{2}\|\boldsymbol{V}_L(\boldsymbol{x}^k) - \boldsymbol{x}^k\|^2.$$

From this inequality, we can see that if $L \geq \beta L_\Psi$, then:

$$F(\boldsymbol{V}_L(\boldsymbol{x}^k)) \leq \mathcal{Q}_L(\boldsymbol{V}_L(\boldsymbol{x}^k); \boldsymbol{x}^k)$$

Hence, Step 4 of Algorithm 1 terminates after a finite number of iterations. □

Let $\mathcal{L}_F(\alpha) = \{x \in \mathcal{X} \mid F(x) \leq \alpha\}$ be the level set of F at α. Now, we are ready to state the following theorem on global convergence of Algorithm 1.

Theorem 2.A.1. *Let $\{\boldsymbol{x}^k\}$ be the sequence generated by Algorithm 1. Then $\{\boldsymbol{x}^k\} \subset \mathcal{L}_F(F(\boldsymbol{x}^0))$ and*

$$\min_{0 \leq k \leq K} \|\boldsymbol{G}_{\beta L_\Psi}(\boldsymbol{x}^k)\|^2 \leq \frac{2(\beta L_\Psi)^2}{L_{\min}(K+1)}\left[F(\boldsymbol{x}^0) - F^\star\right], \tag{2.23}$$

where $F^\star := \inf_{\boldsymbol{x} \in \mathcal{X}} F(\boldsymbol{x}) > -\infty$. Moreover, we also obtain

$$\lim_{k \to \infty} \|\boldsymbol{x}^{k+1} - \boldsymbol{x}^k\| = 0, \quad \textit{and} \quad \lim_{k \to \infty} \|\boldsymbol{G}_{\beta L_\Psi}(\boldsymbol{x}^k)\| = 0, \tag{2.24}$$

and the set of limit points $\hat{\mathcal{S}}^*$ of the sequence $\{\boldsymbol{x}^k\}_{k \geq 0}$ is connected. If this sequence is bounded (in particular, if $\mathcal{L}_F(F(\boldsymbol{x}^0))$ is bounded) then every limit point is a stationary point of (2.2). Moreover, if the set of limit points $\hat{\mathcal{S}}^*$ is finite, then the sequence $\{\boldsymbol{x}^k\}$ converges to a stationary point $\boldsymbol{x}^* \in \mathcal{S}^*$ of (2.2). If, in addition, $\boldsymbol{x}^*$ is feasible to (2.1) and β is sufficiently large, then $\boldsymbol{x}^*$ is also a stationary point of (2.1).

Proof. From Step 5 of Algorithm 1, we have $\boldsymbol{x}^{k+1} := \boldsymbol{V}_{L_k}(\boldsymbol{x}^k)$ and $\{\boldsymbol{x}^k\} \subset \mathcal{X}$. Using (2.20), it is easy to obtain $-\infty < F^\star \leq F(\boldsymbol{x}^{k+1}) \leq F(\boldsymbol{x}^k) \leq \cdots \leq F(\boldsymbol{x}^0)$. This shows that $\{\boldsymbol{x}^k\} \subset \mathcal{L}_F(F(\boldsymbol{x}^0))$, and $\{F(\boldsymbol{x}^k)\}$ is a decreasing sequence and bounded. Hence, it has at least a convergent subsequence. Moreover, from (2.20), we also have

$$F(\boldsymbol{x}^{k+1}) \leq F(\boldsymbol{x}^k) - \frac{L_{\min}}{2} r^2_{L_k}(\boldsymbol{x}^k) \leq F(\boldsymbol{x}^k) - \frac{L_{\min}}{2} r^2_{\beta L_\Psi}(\boldsymbol{x}^k). \tag{2.25}$$

Summing up the inequality (2.25) from $k = 0$ to $k = K$ and using $F(\boldsymbol{x}^{k+1}) \geq F^\star$, we obtain

$$\begin{aligned}\frac{L_{\min}}{2(\beta L_\Psi)^2}\sum_{k=0}^{K}\|\boldsymbol{G}_{\beta L_\Psi}(\boldsymbol{x}^k)\|^2 &= \frac{L_{\min}}{2}\sum_{k=0}^{K} r^2_{\beta L_\Psi}(\boldsymbol{x}^k)\\ &\leq F(\boldsymbol{x}^0) - F(\boldsymbol{x}^{k+1})\\ &\leq F(\boldsymbol{x}^0) - F^\star.\end{aligned}$$

This implies

$$\min_{0\leq k\leq K}\|\boldsymbol{G}_{\beta L_\Psi}(\boldsymbol{x}^k)\|^2 \leq \frac{2(\beta L_\Psi)^2}{L_{\min}(K+1)}\left[F(\boldsymbol{x}^0) - F^\star\right],$$

which leads to (2.23).
Similarly, for any $N \geq 0$ one has

$$F(\boldsymbol{x}^k) - F(\boldsymbol{x}^{k+N}) \geq \frac{L_{\min}}{2}\sum_{i=k}^{k+N-1} r^2_{L_k}(\boldsymbol{x}^i) \geq \frac{L_{\min}}{2}\sum_{i=k}^{k+N-1} r^2_{\beta L_\Psi}(\boldsymbol{x}^i). \tag{2.26}$$

Note that the sequence $\left\{F(\boldsymbol{x}^k)\right\}_{k\geq 0}$ has a convergent subsequence, thus passing to the limit as $k \to \infty$ in (2.26) we obtain the first limit of (2.24). Since $\|\boldsymbol{x}^{k+1} - \boldsymbol{x}^k\| = r_{L_k}(\boldsymbol{x}^k) \geq r_{\beta L_\Psi}(\boldsymbol{x}^k) = \frac{1}{\beta L_\Psi}\|\boldsymbol{G}_{\beta L_\Psi}(\boldsymbol{x}^k)\|$ due to Statement (b) of Lemma 2.A.1, the first limit of (2.24) also implies the second one. If the sequence $\left\{\boldsymbol{x}^k\right\}_{k\geq 0}$ is bounded, by passing to the limit through a subsequence and combining with Lemma 2.A.1, we easily prove that every limit point is a stationary point of (2.2). If the set of limit points $\hat{\mathcal{S}}^*$ is finite, by applying the result in [Ost16][Chapt. 28], we obtain the proof of the remaining conclusion. □

Theorem 2.A.1 provides a global convergence result for Algorithm 1. Moreover, our algorithm requires solving convex subproblems at each iteration, thus offering a great advantage over classical penalty-type schemes. Since the underlying problem is non-convex, the iterates of our algorithm may get trapped at points that may be infeasible for the original problem. That is, under the stated conditions, the iterate sequence $\left\{\boldsymbol{x}^k\right\}$ may converge to a local minimum (stationary) point $\boldsymbol{x}^*$ of (2.2). Since $\boldsymbol{x}^* \in \mathcal{X}$, if $\Psi(\boldsymbol{x}^*) = 0$, then $\boldsymbol{x}^*$ is also a local minimum (stationary) point of the original problem (2.1).

We can sometimes overcome this by combining the algorithm with a *run-and-inspect procedure* [CSY19], whereby if $\boldsymbol{x}^*$ violates $\Psi(\boldsymbol{x}) = 0$, then we restart the algorithm at a new starting point. More precisely, we add an *inspect* phase to our existing algorithm that helps escape from non-feasible stationary points. In the inspection phase, if $\Psi(\boldsymbol{x}^*) \neq 0$ we sample a point around the current

point and increase the parameter β. Since we do not know any optimal Lagrange multiplier $\boldsymbol{y}^*$ of (2.1), we cannot guarantee that $\beta > \|\boldsymbol{y}^*\|_\infty$. However, since it is expected that the multiplier $\boldsymbol{y}_k$ of the subproblem (2.5) converges to $\boldsymbol{y}^*$, we can use $\boldsymbol{y}_k$ to monitor the update of β by guaranteeing that $\beta > |\boldsymbol{y}_k|_\infty$. We have seen that such strategy performs well on a set of realistic non-convex AC-OPF problems. Nevertheless, we do not have theoretical guarantee for this variant and leave this extension for future work.

2.A.2 Local Convergence Analysis

Let us study a special case of (2.2) where Algorithm 1 has a local quadratic convergence rate. Our result relies on the following assumptions. First, for simplicity of our analysis, we assume that $\beta > 0$ is fixed and L_k is also fixed at $L_k := L > 0$ for $k \geq 0$ in Algorithm 1. Next, let $\boldsymbol{x}^*$ be a stationary point of (2.2) such that

$$\langle \nabla f(\boldsymbol{x}^*), \boldsymbol{x} - \boldsymbol{x}^* \rangle + \beta |\Psi'(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*)| \geq \omega_{\min} \|\boldsymbol{x} - \boldsymbol{x}^*\|, \quad \forall \boldsymbol{x} \in \mathcal{N}(\boldsymbol{x}^*) \cap \mathcal{X}, \tag{2.27}$$

where $\omega_{\min} > 0$ is a given constant independent of x and $\mathcal{N}(\boldsymbol{x}^*)$ is a neighborhood of $\boldsymbol{x}^*$. The condition (2.27) is rather technical, but it holds in the following case. Let us assume that the Jacobian $\Psi'(\boldsymbol{x}^*)$ of Ψ at $\boldsymbol{x}^*$ satisfies the following condition

$$\|\Psi'(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*\| \geq \sigma_{\min}(\Psi'(\boldsymbol{x}^*)) \|\boldsymbol{x} - \boldsymbol{x}^*\| \quad \forall\ \boldsymbol{x} \in \mathcal{N}(\boldsymbol{x}^*) \cap \mathcal{X},$$

where $\sigma_{\min}(\Psi'(\boldsymbol{x}^*)$ is the positive smallest singular value of $\Psi'(\boldsymbol{x}^*)$. This condition is similar to the strong second-order sufficient optimality condition [NW06], but only limited to the linear objective function. In this case, we have $|\Psi'(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*)| \geq \|\Psi'(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*)\| \geq \sigma_{\min}(\Psi'(\boldsymbol{x}^*)) \|\boldsymbol{x} - \boldsymbol{x}^*\|$. Therefore, it leads to

$$\langle \nabla f(\boldsymbol{x}^*), \boldsymbol{x} - \boldsymbol{x}^* \rangle + \beta |\Psi'(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*)| \geq (\beta \sigma_{\min}(\Psi'(\boldsymbol{x}^*)) - \|\nabla f(\boldsymbol{x}^*)\|) \|\boldsymbol{x} - \boldsymbol{x}^*\|.$$

For $\beta > 0$ sufficiently large such that $\beta > \frac{\|\nabla f(\boldsymbol{x}^*)\|}{\sigma_{\min}(\Psi'(\boldsymbol{x}^*))}$, we have $\omega_{\min} := \beta \sigma_{\min}(\Psi'(\boldsymbol{x}^*)) - \|\nabla f(\boldsymbol{x}^*)\| > 0$, and the condition (2.27) holds. Now, we prove a local quadratic convergence of Algorithm 1 under assumption (2.27).

Note that a fast local convergence rate such as superlinear or quadratic is usually expected in Gauss-Newton methods, see, e.g., [Kel99, Theorem 2.4.1.]. The following theorem shows that Algorithm 1 can also achieve a fast local quadratic convergence rate under a more restrictive condition (2.27).

Theorem 2.A.2. *Let $\{\boldsymbol{x}_k\}$ be the sequence generated by Algorithm 1 such that it converges to a feasible stationary point $\boldsymbol{x}^*$ of (2.2). Assume further that $\boldsymbol{x}^*$*

satisfies condition (2.27) *for some* $\omega_{\min} > 0$. *Then, if* $\boldsymbol{x}_k \in \mathcal{L}_F(F(\boldsymbol{x}_0))$ *such that* $\|\boldsymbol{x}_k - \boldsymbol{x}^*\| \leq \frac{2\omega_{\min}}{L+(\sqrt{n}+5)\beta L_\Psi}$, *then* $\boldsymbol{x}_{k+1} \in \mathcal{L}_F(F(\boldsymbol{x}_0))$ *and*

$$\|\boldsymbol{x}_{k+1} - \boldsymbol{x}^*\| \leq \left[\frac{L + (\sqrt{n}+3)\beta L_\Psi}{2(\omega_{\min} - L_\Psi\|\boldsymbol{x}_k - \boldsymbol{x}^*\|)}\right]\|\boldsymbol{x}_k - \boldsymbol{x}^*\|^2. \tag{2.28}$$

As a consequence, the sequence $\{\boldsymbol{x}_k\}$ *locally converges to* $\boldsymbol{x}^*$ *at a quadratic rate.*

Proof. Let $d(\boldsymbol{x}, \boldsymbol{y}) := \Psi(\boldsymbol{y}) - \Psi(\boldsymbol{x}) - \Psi'(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x})$. First, by the Lipschitz continuity of $\Psi'(\cdot)$, we have $|d(\boldsymbol{x}, \boldsymbol{y})| \leq \sqrt{n}\|d(\boldsymbol{x}, \boldsymbol{y})\| \leq \frac{\sqrt{n}L_\Psi}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2$. In this case, from this estimate and (2.6), for any $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{L}_F(F(\boldsymbol{x}_0))$, we can derive that

$$\begin{aligned}
\mathcal{Q}_L(\boldsymbol{V}_L(\boldsymbol{x}); \boldsymbol{x}) \quad &:= \min_{\boldsymbol{y}\in\mathcal{X}}\Big\{\mathcal{Q}_L(\boldsymbol{y}; \boldsymbol{x})\Big\} \\
&= \min_{\boldsymbol{y}\in\mathcal{X}}\Big\{f(\boldsymbol{y}) + \beta|\Psi(\boldsymbol{x}) + \Psi'(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x})| + \frac{L}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2\Big\} \\
&= \min_{\boldsymbol{y}\in\mathcal{X}}\Big\{f(\boldsymbol{y}) + \beta|\Psi(\boldsymbol{y}) - d(\boldsymbol{x}, \boldsymbol{y})| + \frac{L}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2\Big\} \\
&\leq \min_{\boldsymbol{y}\in\mathcal{X}}\Big\{f(\boldsymbol{y}) + \beta|\Psi(\boldsymbol{y}))| + \beta|d(\boldsymbol{x}, \boldsymbol{y})| + \frac{L}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2\Big\} \\
&\leq \min_{\boldsymbol{y}\in\mathcal{X}}\Big\{f(\boldsymbol{y}) + \beta|\Psi(\boldsymbol{y}))| + \frac{(L + \sqrt{n}\beta L_\Psi)}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2\Big\} \\
&\leq \min_{\boldsymbol{y}\in\mathcal{X}}\Big\{F(\boldsymbol{y}) + \frac{(L + \sqrt{n}\beta L_\Psi)}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2\Big\}.
\end{aligned}$$

This estimate together with $\Psi(\boldsymbol{x}^*) = 0$ and $\boldsymbol{x} = \boldsymbol{x}_k$ imply that

$$\begin{aligned}
\mathcal{Q}_L(\boldsymbol{x}_{k+1}; \boldsymbol{x}_k) \equiv \mathcal{Q}_L(\boldsymbol{V}_L(\boldsymbol{x}_k); \boldsymbol{x}_k) \quad &\leq F(\boldsymbol{x}^*) + \frac{(L + \sqrt{n}\beta L_\Psi)}{2}\|\boldsymbol{x}_k - \boldsymbol{x}^*\|^2 \\
&= f^\star + \frac{(L + \sqrt{n}\beta L_\Psi)}{2}\|\boldsymbol{x}_k - \boldsymbol{x}^*\|^2.
\end{aligned}$$

Moreover, by μ_f-convexity of f, we have $f(\boldsymbol{x}_{k+1}) - f^\star \geq \langle\nabla f(\boldsymbol{x}^*), \boldsymbol{x}_{k+1} - \boldsymbol{x}^*\rangle + \frac{\mu_f}{2}\|\boldsymbol{x}_{k+1} - \boldsymbol{x}^*\|^2$. Using these last two estimates and the definition of $\mathcal{Q}_L$, we

can show that

$$\begin{aligned}
&\frac{(L+\sqrt{n}\beta L_\Psi)}{2}\|\boldsymbol{x}_k-\boldsymbol{x}^*\|^2\\
&\quad\geq f(\boldsymbol{x}_{k+1})-f^\star+\beta|\Psi(x_k)+\Psi'(\boldsymbol{x}_k)(\boldsymbol{x}_{k+1}-\boldsymbol{x}_k)|+\frac{L}{2}\|\boldsymbol{x}_{k+1}-\boldsymbol{x}_k\|^2\\
&\quad\geq \frac{\mu_f}{2}\|\boldsymbol{x}_{k+1}-\boldsymbol{x}^*\|^2+\langle\nabla f(\boldsymbol{x}^*),\boldsymbol{x}_{k+1})-\boldsymbol{x}^*\rangle+\beta|\Psi'(\boldsymbol{x}_k)(\boldsymbol{x}_{k+1}-\boldsymbol{x}^*)\\
&\qquad+\Psi(\boldsymbol{x}_k)-\Psi(\boldsymbol{x}^*)-\Psi'(\boldsymbol{x}^*)(\boldsymbol{x}_k-\boldsymbol{x}^*)+(\Psi'(\boldsymbol{x}_k)-\Psi'(\boldsymbol{x}^*))(\boldsymbol{x}_{k+1}-\boldsymbol{x}_k)|\\
&\quad\geq \frac{\mu_f}{2}\|\boldsymbol{x}_{k+1}-\boldsymbol{x}^*\|^2+\langle\nabla f(\boldsymbol{x}^*),\boldsymbol{x}_{k+1})-\boldsymbol{x}^*\rangle+\beta|\Psi'(\boldsymbol{x}^*)(\boldsymbol{x}_{k+1}-\boldsymbol{x}^*)|\\
&\qquad-\frac{\beta L_\Psi}{2}\|\boldsymbol{x}_k-\boldsymbol{x}^*\|^2-\beta L_\Psi\|\boldsymbol{x}_k-\boldsymbol{x}^*\|\|\boldsymbol{x}_{k+1}-\boldsymbol{x}_k\|.
\end{aligned}$$

Finally, using condition (2.27), we obtain from the last estimate that

$$\begin{aligned}
&\frac{(L+\sqrt{n}\beta L_\Psi)}{2}\|\boldsymbol{x}_k-\boldsymbol{x}^*\|^2\\
&\qquad\geq(\omega_{\min}-\beta L_\Psi\|\boldsymbol{x}_k-\boldsymbol{x}^*\|)\,\|\boldsymbol{x}_{k+1}-\boldsymbol{x}^*\|-\frac{3\beta L_\Psi}{2}\|\boldsymbol{x}_k-\boldsymbol{x}^*\|^2.
\end{aligned}$$

Rearranging this inequality given that $\|\boldsymbol{x}_k-\boldsymbol{x}^*\|<\frac{\omega_{\min}}{\beta L_\Psi}$, we obtain (2.28). From (2.28), we can see that if $\|\boldsymbol{x}_k-\boldsymbol{x}^*\|\leq\frac{2\omega_{\min}}{L+(\sqrt{n}+5)\beta L_\Psi}$, then $\|\boldsymbol{x}_k-\boldsymbol{x}^*\|<\frac{\omega_{\min}}{\beta L_\Psi}$. Moreover, $\|\boldsymbol{x}_{k+1}-\boldsymbol{x}^*\|\leq\|\boldsymbol{x}_k-\boldsymbol{x}^*\|$. Hence, if $\boldsymbol{x}_k\in\mathcal{L}_F(F(\boldsymbol{x}_0))$, then $\boldsymbol{x}_{k+1}\in\mathcal{L}_F(F(\boldsymbol{x}_0))$. The last statement is a direct consequence of (2.28). □

2.B Parameter Tuning of Algorithm 1

In order to explain how the parameters of the GN algorithm should be tuned, we first express the subproblem at iteration k for the special case of AC-OPF.

Subproblem (2.6) **in the context of AC-OPF.** We define the following functions in order to keep our notation compact:

$$\Phi_q(\boldsymbol{c},\boldsymbol{s},\boldsymbol{dc},\boldsymbol{ds})=\sum_{(i,j)\in\mathcal{L}}\left|\Psi_q^{ij}(\boldsymbol{c},\boldsymbol{s})+\Psi_q^{ij\,\prime}(\boldsymbol{c},\boldsymbol{s})(\boldsymbol{dc},\boldsymbol{ds})^\top\right|,$$

$$\Phi_t(\boldsymbol{c},\boldsymbol{s},\boldsymbol{\theta},\boldsymbol{dc},\boldsymbol{ds},\boldsymbol{d\theta})=\sum_{(i,j)\in\mathcal{L}}\left|\Psi_t^{ij}(\boldsymbol{c},\boldsymbol{s},\boldsymbol{\theta})+\Psi_t^{ij\,\prime}(\boldsymbol{c},\boldsymbol{s},\boldsymbol{\theta})(\boldsymbol{dc},\boldsymbol{ds},\boldsymbol{d\theta})^\top\right|.$$

Note that we are facing two types of constraints (*quadratic* and *trigonometric*) and that they may attain different relative scales for different instances, as we have observed in our numerical experiments. We will define different penalty terms depending on the type of constraint, which will increase the flexibility

of our implementation, while respecting the theoretical assumptions. To this end, we denote by β_q (resp. β_t) the β penalty parameter associated with the quadratic (resp. trigonometric) constraints. The trigonometric constraints depend on $\boldsymbol{c}$, $\boldsymbol{s}$ and $\boldsymbol{\theta}$, whereas the quadratic ones only depend on $\boldsymbol{c}$ and $\boldsymbol{s}$. Therefore, we define two separate regularization parameters: L_{cs} and L_θ (we drop the k index from now on for notational simplicity).

The GN subproblem at iteration k, corresponding to (2.5), is convex and has the form:

$$\begin{aligned}\mathcal{P}^k_{sub}: \quad \min_{(\boldsymbol{y},\boldsymbol{d})} \quad & f(\boldsymbol{x}) + \beta_q \Phi_q(\boldsymbol{c}^k, \boldsymbol{s}^k, \boldsymbol{dc}, \boldsymbol{ds}) + \beta_t \Phi_t(\boldsymbol{c}^k, \boldsymbol{s}^k, \boldsymbol{\theta}^k, \boldsymbol{dc}, \boldsymbol{ds}, \boldsymbol{d\theta}) \\ & + \frac{L_{cs}}{2}\|(\boldsymbol{dc}, \boldsymbol{ds})\|^2 + \frac{L_\theta}{2}\|\boldsymbol{d\theta}\|^2 \\ s.t. \quad & \boldsymbol{y} = (\boldsymbol{p}, \boldsymbol{q}),\ \boldsymbol{d} = (\boldsymbol{dc}, \boldsymbol{ds}, \boldsymbol{d\theta}) \\ & (\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{c}^k + \boldsymbol{dc}, \boldsymbol{s}^k + \boldsymbol{ds}, \boldsymbol{\theta}^k + \boldsymbol{d\theta}) \in \mathcal{X}.\end{aligned}$$

The optimal solution of $\mathcal{P}^k_{sub}$, which we denote by $(\boldsymbol{y}^\star, \boldsymbol{d}^\star)$, provides the next iterate $\boldsymbol{x}^{k+1} = (\boldsymbol{p}^\star, \boldsymbol{q}^\star, \boldsymbol{c}^k + \boldsymbol{dc}^\star, \boldsymbol{s}^k + \boldsymbol{ds}^\star, \boldsymbol{\theta}^k + \boldsymbol{d\theta}^\star)$.

2.B.1 Joint Values of Parameters: $\beta_q = \beta_t$ and $L_{cs} = L_\theta$

In our first set of tests, we implement the basic variant of Algorithm 1 by retaining the configuration of parameters from our theoretical results. Since we only aim at validating the algorithm, we focus on the three instances that have less than 2,000 nodes: `1354pegase`, `1888rte`, and `1951pegase`.

For the penalty parameters, we first choose the same value β for both quadratic and trigonometric constraints as $\beta = \beta^q = \beta^t$. We perform tests with three choices of β, based on the number of lines in the test case: $\beta = |\mathcal{L}|/100, |\mathcal{L}|/10$, or $|\mathcal{L}|$. For the regularization parameter L, we also choose the same value $L = L_{cs} = L_\theta$. Note that all the experiments are initialized from the SOCP solution. We consider three different strategies when applying Algorithm 1:

- *Fixed strategy*: we fix L at the upper bound L_Ψ, which is computed in (2.10).
- *Bisection update*: At each iteration k of Algorithm 1, we choose $L_{\min} = 1$ and initialize $L := 1$. If we do not satisfy the line-search condition $F(V_L(\boldsymbol{x}^k)) \leq \mathcal{Q}_L(V_L(\boldsymbol{x}^k); \boldsymbol{x}^k)$, we apply a bisection in the interval $[L_{\min}, \beta L_\psi]$ until we satisfy this condition.
- *Geometric-μ update*: At each iteration k, we also choose $L_{\min} = 1$ and initialize $L := 1$. We then update $L \leftarrow \mu \cdot L$ for $\mu > 1$ in order to guarantee that $F(V_L(\boldsymbol{x}^k)) \leq \mathcal{Q}_L(V_L(\boldsymbol{x}^k); \boldsymbol{x}^k)$. In our experiments, we use $\mu := 2$.

The results of our first test are presented in Table 2.10. The three columns for each strategy present the number of iterations, and the maximum violation of the quadratic (resp. trigonometric) constraint for each strategy on the three problem instances. In terms of number of iterations, we can observe that fixing L results in a poor performance and tends to increase the number of iterations as well as the final violation of the constraints. Given this poor performance, we do not consider the *Fixed* strategy for the remainder of the numerical experiments. Also, using $\beta = |\mathcal{L}|$ produces satisfactory results, and increasing the value of β supports convergence, therefore we choose this order of magnitude for the value of β.

Table 2.10: Performance results of the Gauss-Newton algorithm with different strategies for β and L.

		Fixed			Bisection			Geometric-2		
Test Case	β	# It	MVQ	MVT	# It	MVQ	MVT	# It	MVQ	MVT
`1354pegase`	$\|\mathcal{L}\|/100$	100	$6e^{-3}$	$4e^{-6}$	100	$6e^{-3}$	$4e^{-6}$	100	$7e^{-3}$	$4e^{-6}$
	$\|\mathcal{L}\|/10$	91	$2e^{-6}$	$6e^{-9}$	3	$5e^{-6}$	$2e^{-7}$	3	$5e^{-6}$	$4e^{-6}$
	$\|\mathcal{L}\|$	78	$2e^{-6}$	$1e^{-8}$	3	$8e^{-9}$	$4e^{-9}$	3	$6e^{-6}$	$5e^{-6}$
`1888rte`	$\|\mathcal{L}\|/100$	100	$1e^{-2}$	$8e^{-8}$	21	$1e^{-2}$	$9e^{-8}$	100	$7e^{-3}$	$8e^{-7}$
	$\|\mathcal{L}\|/10$	100	$8e^{-3}$	$5e^{-8}$	11	$7e^{-3}$	$7e^{-8}$	100	$5e^{-3}$	$3e^{-8}$
	$\|\mathcal{L}\|$	100	$8e^{-3}$	$5e^{-8}$	15	$5e^{-3}$	$2e^{-8}$	100	$3e^{-3}$	$8e^{-8}$
`1951rte`	$\|\mathcal{L}\|/100$	86	$4e^{-6}$	$4e^{-9}$	10	$2e^{-3}$	$5e^{-8}$	39	$3e^{-4}$	$4e^{-8}$
	$\|\mathcal{L}\|/10$	66	$6e^{-7}$	$3e^{-8}$	4	$2e^{-5}$	$6e^{-8}$	3	$8e^{-6}$	$6e^{-6}$
	$\|\mathcal{L}\|$	61	$1e^{-6}$	$1e^{-7}$	4	$5e^{-7}$	$1e^{-7}$	3	$3e^{-6}$	$2e^{-6}$

Bisection and *Geometric-2* exhibit a similar behavior: when the algorithm converges to a feasible point, both choices achieve converge in tens of iterations, depending on the test case and the choice of β. Nevertheless, when failing, the maximum violation of the quadratic constraint (MVQ) never reaches the desired tolerance of $1e^{-5}$. This behavior might suggest that we do not penalize sufficiently the quadratic constraint. One should notice that the quadratic and trigonometric constraints are linked: once the angles are fixed, $\boldsymbol{c}$ and $\boldsymbol{s}$, which are the variables that appear in the quadratic constraints, struggle to move from their current value in order to satisfy the quadratic constraints. This motivates us to consider two different values for β and L: β_q for the quadratic constraints and L_{cs} for the associated variables $\boldsymbol{c}$ and $\boldsymbol{s}$; β_t for the trigonometric constraints and L_θ for the angle variables $\boldsymbol{\theta}$.

2.B.2 Adaptation and Enhancement of Algorithm 1 for AC-OPF

Adapting Algorithm 1 with Individual Choices of Parameters: β_q, β_t and L_{cs}, L_θ. Based on our observations, we choose different values for β_q and β_t. Since we empirically observe that the quadratic constraints are harder

to satisfy than the trigonometric constraints, we consider two alternatives: $\beta_q = 2\beta_t$ and $\beta_q = 5\beta_t$. Furthermore, we set $L_{cs,\min} = \beta_q/\beta_t$ and $L_{\theta,\min} = 1$. Note that the choice of different values for these parameters does not affect the theoretical guarantees of our algorithm, as long as they satisfy our given conditions.

Also, since we have different values L_{cs} and L_θ, we must adapt the condition under which L_{cs} and L_θ are updated. From the theory, L is updated (through the *Bisection* or *Geometric* strategy) if the following condition is not met:

$$\begin{aligned} & F(\boldsymbol{x}^{k+1}) \leq \mathcal{Q}_L(\boldsymbol{x}^{k+1}; \boldsymbol{x}^k) \\ \Leftrightarrow \quad & \beta_q \sum_{(i,j)\in\mathcal{L}} |\Psi_q^{ij}(\boldsymbol{c}^{k+1}, \boldsymbol{s}^{k+1})| + \beta_t \sum_{(i,j)\in\mathcal{L}} |\Psi_t^{ij}(\boldsymbol{c}^{k+1}, \boldsymbol{s}^{k+1}, \boldsymbol{\theta}^{k+1})| \\ & \leq \beta_q \phi_q(\boldsymbol{c}^k, \boldsymbol{s}^k, \boldsymbol{dc}^*, \boldsymbol{ds}^*) + \beta_t \phi_t(\boldsymbol{c}^k, \boldsymbol{s}^k, \boldsymbol{\theta}^k, \boldsymbol{dc}^*, \boldsymbol{ds}^*, \boldsymbol{d\theta}^*) \\ & \quad + \frac{L_{cs}}{2}\|(\boldsymbol{dc}^*, \boldsymbol{ds}^*)\|^2 + \frac{L_\theta}{2}\|(\boldsymbol{d\theta}^*)\|^2 \end{aligned}$$

where $\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \boldsymbol{d}^*$.

We adapt this condition to the specific type of constraint. Concretely:

- If

$$\begin{aligned} & \beta_t \sum_{(i,j)\in\mathcal{L}} |\Psi_t^{ij}(\boldsymbol{c}^{k+1}, \boldsymbol{s}^{k+1}, \boldsymbol{\theta}^{k+1})| \\ & \leq \beta_t \Phi_t(\boldsymbol{c}^k, \boldsymbol{s}^k, \boldsymbol{\theta}^k, \boldsymbol{dc}^*, \boldsymbol{ds}^*, \boldsymbol{d\theta}^*) + \frac{L_{cs}}{2}||(\boldsymbol{dc}^*, \boldsymbol{ds}^*)^\top||^2 + \frac{L_\theta}{2}||(\boldsymbol{d\theta}^*)||^2 \end{aligned} \tag{2.29}$$

  then update L_{cs} and L_θ.

- If (2.29) does not hold and

$$\beta_q \sum_{(i,j)\in\mathcal{L}} |\Psi_q^{ij}(\boldsymbol{c}^{k+1}, \boldsymbol{s}^{k+1})| \leq \beta_q \Phi_q(\boldsymbol{c}^k, \boldsymbol{s}^k, \boldsymbol{dc}^*, \boldsymbol{ds}^*) + \frac{L_{cs}}{2}||(\boldsymbol{dc}, \boldsymbol{ds})^\top||^2 \tag{2.30}$$

  then only update L_{cs}.

Refining the L updates. We emphasize that, in Algorithm 1, each time that the values of L_{cs} and/or L_θ are updated, the subproblem $\mathcal{P}_{sub}^k$ is resolved. Therefore, an effective strategy for updating these parameters can lead to significant improvements in computational time. We mitigate this heavy computational requirement by introducing resolution techniques that are guided by both our theoretical results and empirical observations. Concretely, we propose the following two improvements to the practical implementation of the algorithm, in order to limit the number of computationally expensive L updates:

1. Keeping the value of L from one iteration to another is a better strategy than having $L := L_{\min}$ at the beginning of each iteration. The natural justification for this is that we expect steps to decrease along the iterations.

2. Checking conditions (2.29) and (2.30) can require a large number of L updates. Instead, we propose a verification of whether the violation of the constraints is decreasing before checking (2.29) and (2.30), which is a less stringent requirement that still yields satisfactory results in terms of constraint violations. Concretely, at iteration k, we compute the ℓ_1 and ℓ_∞ norms of $\Psi_q(\boldsymbol{c}^k, \boldsymbol{s}^k)$ and $\Psi_t(\boldsymbol{c}^k, \boldsymbol{s}^k, \boldsymbol{\theta}^k)$. If these quantities decrease from $k-1$ to k, we move to iteration $k+1$.

By applying these two approaches, we obtain the results that are presented in Table 2.11. In Table 2.11, # It provides the number of GN iterations, # L-It provides the number of additional subproblems solved because of an L-update (then # It+# L-It gives the total number of subproblems solved) and Obj records the objective value returned. From Table 2.11, we observe that this new strategy leads to convergence for all three test cases. Even if Bisection seems to require less iterations, it also provides higher objective values than Geometric-2. Also, we do not observe notable differences between applying a factor of 2 or 5 on β, although a factor of 5 decreases slightly the number of iterations. In our implementation, we employ the following settings: Geometric-2, $\beta_t = |\mathcal{L}|$ and $\beta_q = 5\beta_t$.

Table 2.11: Performance behavior of the Gauss-Newton algorithm with different values and strategies for β_q, β_t, L_{cs}, L_θ.

			Bisection			Geometric-2		
Test Case	β_t	β_q	# It	# L-It	Obj	# It	# L-It	Obj
`1354pegase`	$\|\mathcal{L}\|$	$\times 2$	3	1	7.408e4	14	6	7.407e4
	$\|\mathcal{L}\|$	$\times 5$	4	1	7.408e4	13	4	7.407e4
`1888rte`	$\|\mathcal{L}\|$	$\times 2$	8	1	5.982e4	17	1	5.981e4
	$\|\mathcal{L}\|$	$\times 5$	11	1	5.982e4	14	3	5.981e4
`1951rte`	$\|\mathcal{L}\|$	$\times 2$	4	1	8.174e4	6	1	8.174e4
	$\|\mathcal{L}\|$	$\times 5$	4	0	8.174e4	4	0	8.174e4

Chapter 3

Coordination Schemes for the Integration of T&D System Operations

3.1 Introduction

The mobilization of distributed resources is emerging as an increasingly important and challenging aspect of power system operations [PAK16]. This paradigm shift towards the proactive management of distributed resources is driven by a number of factors that are influencing the transition of the energy industry, including: (i) the large-scale integration of renewable resources is placing ever-increasing needs on power system flexibility; (ii) residential and commercial demand, which is connected to low-voltage grids, represents the majority of demand-side flexibility [Gil14]; (iii) the proliferation of distributed renewable supply and distributed storage requires an intelligent management of distribution power flows in order to postpone or avoid costly distribution network infrastructure upgrades; (iv) communication and control technology appear to offer adequate technological solutions to the computation, communication and control requirements of this transition.

The integration of distributed resources in proactive power system operations poses two major challenges from a modeling and computational perspective: the number of resources is vast, and the physics of distribution networks cannot be adequately represented through linearized power flow models. A direct approach towards the integration of distributed resources involves integrated optimization whereby the transmission system and the distribution system are optimized simultaneously was proposed recently by Caramanis et al. [CNH+16], which exploits recent breakthroughs on conic relaxations of opti-

mal power flow [FL13]. The distribution locational marginal price signals generated from this approach endogenize the value of losses, voltage constraints, and complex power flow constraints. In such a scheme, the operations of the distribution system operator (DSO) are effectively absorbed by the transmission system operator (TSO). Hierarchical approaches whereby the DSO reacts to a locational marginal price at the TSO-DSO interface have also been suggested in the literature [SV06, DPN$^+$13, LWO14, VVL14, HWO$^+$15]. Such hierarchical schemes may be required in order to achieve scalability in the mobilization of distributed resources. However, the detailed description and modeling of TSO-DSO coordination has yet to be clarified in the literature. A particularly challenging aspect of TSO-DSO coordination is the extent to which system imbalances can be resolved through the mobilization of distributed reserves while respecting distribution system constraints, and which entity, the TSO or the DSO, should be responsible for this decision.

The SmartNet consortium has proposed various coordination schemes for TSO and DSO operations that aim at the scalable mobilization of distributed resources [GRS16]. The focus of SmartNet is on the *activation* of reserve capacity. The contribution of this chapter is to propose models that can be used for the quantitative evaluation of TSO-DSO coordination schemes. The models presented in this chapter can be used for assessing the allocative efficiency, the 'proximity' of dispatch to physically compatible solutions and the price signals generated by various TSO-DSO coordination schemes. This chapter also serves as a basis for chapters 4 and 5. Table 3.1 gives an overview of the coordination schemes and their use throughout the rest of the dissertation.

The remainder of the chapter is organized as follows. Section 3.2 presents models for the SmartNet coordination schemes. Section 3.3 demonstrates the proposed models on a small-scale test system. Section 3.4 concludes the chapter.

3.2 Modeling TSO-DSO Coordination Schemes

The following models consider the activation of reserves for balancing a real-time deviation in the net load of a transmission or distribution bus. It is assumed that reserve capacity and a predetermined dispatch have been cleared in earlier markets, and reserve activation bids are offered for upward activation or downward activation in real time.

Five schemes will be considered: (i) Centralized Common TSO-DSO market, (ii) Decentralized Common TSO-DSO market, (iii) Centralized Ancillary Services Market, (iv) Local Ancillary Services Market, and (v) Shared Balancing Responsibility. Each of these schemes is described in detail, and for each of the schemes a model is proposed in this section. The topology of the networks considered in this chapter consists of meshed transmission networks and

Table 3.1: T&D Coordination schemes considered and their application in the dissertation.

Scheme	Principle	Explored further?
Centralized Common Market	Considers a single operator handling the complete T&D network. Unrealistic in practice but serves as a benchmark.	Serves as benchmark in chapters 4 and 5.
Decentralized Common Market	Consists of maximizing the global welfare while only sharing interface/border information between the TSO and the DSO. The interface flow is priced using a residual supply function by the DSO and cleared by the TSO.	Implemented and detailed in chapter 4 under the name 'RSF Approach'.
Centralized Ancillary Services Market	Prevalent paradigm in which the distribution imbalance is aggregated at the transmission level to clear the wholesale market.	Leads to infeasible dispatch because of the omission of losses when aggregating. Not further explored.
Local Ancillary Services Market	The DSO solves local imbalances and provides the remaining capacity available to the TSO. The interface flow is decided by the TSO.	Not straightforward to model. Leads to infeasible dispatch in chapter 3 due to the difficulty to anticipate losses. Two alternatives studied in chapter 5 (TLA and LM) and assessed using game theory.
Shared Balancing Responsibilities	No communication between the TSO and the DSO. Each operator in charge of its own network. The interface flow is fixed to a predetermined value neither decided by the TSO nor the DSO.	Feasible dispatches with potentially excessive cost in certain configurations. Assessed using game theory in chapter 5.

radial distribution networks, as indicated in Figure 3.2. For simplicity, we will write the models with one transmission network and one distribution network (Figure 3.1). Extending these models to several distribution networks is trivial.

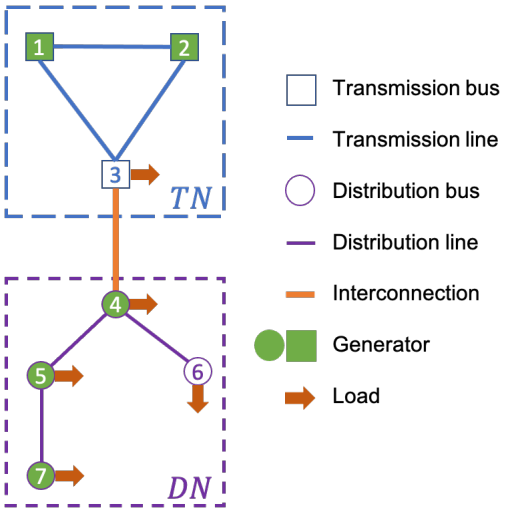


Figure 3.1: Illustration of the T&D network structure considered in Chapters 3, 4 and 5.

3.2.1 Centralized Common TSO-DSO Market

This approach is based on Caramanis et al. [CNH+16]. Transmission and distribution network resources are dispatched according to an integrated optimization of the entire system. The goal of the system operator is to minimize the cost of reserve activation.

The transmission network is represented through linearized DC power flow approximation (similar to (1.7)):

$$F_i^{DC}(x^T) = 0, \qquad \forall i \in \mathcal{TB} \tag{3.1a}$$
$$(\theta, f^p_{\mathcal{TL}}) \in \mathcal{OC}^{DC} \tag{3.1b}$$
$$p_g \in \mathcal{GC}_g^{DC}, \qquad \forall g \in \mathcal{TG} \tag{3.1c}$$

The distribution network is represented through a second-order cone programming (SOCP) relaxation, which is tight for radial distribution networks under mild assumptions [FL13,GLTL14]. The formulation is the same as (1.14).

$$F_i^{SOC}(x^D) = 0, \qquad \forall i \in \mathcal{DB} \tag{3.2a}$$
$$(c, s, f^p_{\mathcal{DL}}, f^q) \in \mathcal{OC}^{SOC} \tag{3.2b}$$
$$(p_g, q_g) \in \mathcal{GC}_g, \qquad \forall g \in \mathcal{DG} \tag{3.2c}$$

In addition, transmission and distribution networks are linked through the interconnection $l' = (i', j')$. We assume that this line is a transmission line only

modeled by $f^p_{i'j'}$ (for simplicity we drop the angles $\theta_{i'},\ \theta_{j'}$). For ease of modeling, we will also assume that variable $f^p_{i'j'}$ is duplicated in the transmission and in the distribution network. By defining $\mathcal{IOC} = \{f^p_{i'j'} | \underline{f}^p_{i'j'} \leq f^p_{i'j'} \leq \overline{f}^p_{i'j'}\}$, the interconnection constraints are the following:

$$f^p_{i'j'} = f^{p,T}_{i'j'} \tag{3.3a}$$
$$f^p_{i'j'} = f^{p,D}_{i'j'} \tag{3.3b}$$
$$f^p_{i'j'} \in \mathcal{IOC} \tag{3.3c}$$

The Centralized Common TSO-DSO aims at minimizing the reserve activation costs. Denoting by $\tilde{p}_g$ a predetermined dispatch and by $r_g \geq 0$ the reserve capacity cleared for generator $g \in \mathcal{G}$, the decision in real-time can be seen as an adjustment Δp_g on the dispatch $\tilde{p}_g$. In a balancing model, we then typically have the following capacity constraints:

$$\tilde{p}_g - r_g \leq \tilde{p}_g + \Delta p_g \leq \tilde{p}_g + r_g$$

By substituting $\underline{p}_g = \tilde{p}_g - r_g$, $p_g = \tilde{p}_g + \Delta p_g$, and $\overline{p}_g = \tilde{p}_g + r_g$, it is possible to cast the balancing problem as an OPF problem as follows:

$$\min_{x^T, x^D, f^p_{i'j'}} \quad \sum_{g \in \mathcal{G}} C_g(p_g) \tag{3.4a}$$
$$s.t. \quad (3.1), (3.2), (3.3) \tag{3.4b}$$
$$x^T = (p_{\mathcal{TG}}, f^p_{\mathcal{TL}}, \theta, f^{p,T}_{i'j'}),$$
$$x^D = (p_{\mathcal{DG}}, q, c, s, f^p_{\mathcal{DL}}, f^q, f^{p,D}_{i'j'})$$

3.2.2 Decentralized Common TSO-DSO Market

This market is modeled by using a *residual* supply function for real power at the balancing market operated by the TSO. This residual supply function is computed by the DSO, which operates its own local market while accounting for its private distribution network constraints [ASL+17]. To be more specific, the DSO computes the residual supply function $TC_{i'j'}(f^p_{i'j'})$ by solving the following problem for different values of $f^p_{i'j'}$:

$$TC_{i'j'}(f^p_{i'j'}) = \min_{x^D} \quad \sum_{g \in \mathcal{DG}} C_g(p_g) \tag{3.5a}$$
$$s.t. \quad (3.2), (3.3c) \tag{3.5b}$$
$$(\mu_{i'j'}) : f^p_{i'j'} = f^{p,D}_{i'j'} \tag{3.5c}$$

Note that this problem needs to be solved *before* the clearing of the TSO balancing market, and is therefore probably agnostic about the actual realization of real-power imbalances in the distribution network. In addition, the communication of a detailed residual function $TC_{i'j'}(\cdot)$ may be excessively onerous in terms of computation and communication, therefore it is assumed that the DSO only communicates a *linearization* of the residual function around a set of predetermined points $\tilde{f}_l^p$:

$$TC_{i'j'}(f^p) \simeq \max_{\tilde{f}_l^p}(TC_{i'j'}(\tilde{f}_l^p) + \mu_l \cdot (f^p - \tilde{f}_l^p)),$$

where $\mu_l \in \partial TC_{i'j'}(\tilde{f}_l^p)$ is a subgradient of $TC_{i'j'}$ at the operating point $\tilde{f}_l^p$. If the market clearing problem of the DSO is convex[1], the function $TC_{i'j'}$ can be seen to be convex. In the case where the distribution network subproblem is not convex (because of block orders and/or if one considers AC modeling), the idea is to rely on the convex hull [GNB20] of the residual supply function. In the scope of the thesis, the empirical testing of chapter 4 shows satisfying results without having to carefully model the convex hull of the residual supply function. The question of approximating a non-convex residual supply function is then left for future research. The requisite data that is needed for the linearization of the function can be obtained by the DSO market clearing problem: $TC_{i'j'}(f^p_{i'j'})$ is simply the objective function of the distribution network subproblem (3.5), while μ_i can be obtained from the dual optimal multiplier of constraint (3.5c), fixing the interface flow.

With the residual supply function (or its linear approximation) in place, the TSO can then solve the following balancing problem:

$$\min_{x^T, f^p_{i'j'}} \sum_{g \in \mathcal{TG}} C_g(p_g) + TC_{i'j'}(f^p_{i'j'}) \tag{3.6a}$$

$$s.t. \quad (3.1), (3.3a), (3.3c) \tag{3.6b}$$

The resulting real power flow at the interconnection is injected to or withdrawn from the DSO system. The distribution network is then dispatched, given the resulting real power flow at the interconnection. This coordination scheme takes advantage of the fact that the only thing that the TSO and the DSO need to agree on is the real power flow at the interface. This hierarchical control should therefore deliver a near-optimal performance, with reasonable communication requirements between TSO and DSO. Note that the financial roles and responsibilities of each entity are also well defined. The DSO participates in the TSO market through an energy bid, and receives a payment from (or pays to) the TSO for its cleared quantity. The DSO then clears its

[1]This can be achieved if we use the SOCP relaxation of the distribution network subproblem.

local market, accounting for its local constraints, and using the previously collected payment from the participation in the TSO auction in order to distribute payments to its local market participants. A similar paradigm has been described in a recent publication by Kristov [KMT16]. Note that this scheme also faces several challenges: (i) the question of the precision of the approximation of $TC()$ can largely influence the efficiency of the scheme, in particular when considering a time-horizon; (ii) each point on which $TC()$ is evaluated implies solving a second-order cone program and the computational burden of such a scheme can be problematic on realistic instances. Chapter 4 provides a complete market-clearing platform embedding this coordination scheme.

3.2.3 Centralized Ancillary Services Market

In this approach, the TSO clears a market for ancillary services at the transmission level, using resources from the transmission and distribution system, but without accounting for distribution network constraints. In order not to violate distribution network constraints, resources need to be pre-qualified, in the sense that distribution resources are not offered in the TSO market if they may violate distribution network constraints. This pre-qualification process is not modeled explicitly in this chapter, instead it is assumed that pre-qualification has already been concluded.

The centralized ancillary services market model dispatches the system so as to relieve the imbalance that has occurred by aggregating distributed resources that offer reserve to their interface line. The TSO solves the following model:

$$\min_{x^T, p_{\mathcal{DG}}, f^p_{i'j'}} \quad \sum_{g \in \mathcal{G}} C_g(p_g) \tag{3.7a}$$

$$s.t. \quad (3.1), (3.3c) \tag{3.7b}$$

$$f^p_{i'j'} = \sum_{g \in \mathcal{DG}} p_g - \sum_{i \in \mathcal{DB}} D^p_i \tag{3.7c}$$

$$\underline{p}_g \leq p_g \leq \overline{p}_g, \qquad \forall g \in \mathcal{DG} \tag{3.7d}$$

Note that this coordination scheme ignores real power losses and shunt capacitance losses.

3.2.4 Local Ancillary Services Market

The local ancillary services market model activates resources depending on where the imbalance occurs. For transmission-level imbalances, the dispatch is

obtained as follows:

$$\min_{x^T, p_{\mathcal{DG}}, f^p_{i'j'}} \quad \sum_{g \in \mathcal{G}} C_g(p_g) \tag{3.8a}$$

$$s.t. \quad (3.1), (3.3c), (3.7c) \tag{3.8b}$$

$$\underline{RLAST}_g \leq p_g \leq \overline{RLAST}_g, \qquad \forall g \in \mathcal{DG} \tag{3.8c}$$

where $\underline{RLAST}_g/\overline{RLAST}_g$ corresponds to the ancillary services capacity that is available for transmission system balancing from generator g of the distribution network.

For a distribution system imbalance, the dispatch is obtained by fixing the amount of real power flow at the interconnection to the value obtained in (3.8), denoted $\tilde{f}^p_{i'j'}$:

$$\min_{x^D, f^p_{i'j'}} \quad \sum_{g \in \mathcal{DG}} C_g(p_g) \tag{3.9a}$$

$$s.t. \quad (3.2) \tag{3.9b}$$

$$f^p_{i'j'} = \tilde{f}^p_{i'j'}, \tag{3.9c}$$

$$\underline{RLASD}_g \leq p_g \leq \overline{RLASD}_g, \qquad \forall g \in \mathcal{DG} \tag{3.9d}$$

where $\underline{RLASD}_g/\overline{RLASD}_g$ corresponds to the ancillary services capacity that is available for distribution system balancing from the distribution network.

3.2.5 Shared Balancing Responsibility

The shared balancing responsibility requires that the TSO clear transmission-level imbalances by using transmission-level resources *only*, and the DSO clear distribution-level imbalances by using distribution-level resources *only*.

The TSO subproblem is modeled as

$$\min_{x^T, f^p_{i'j'}} \quad \sum_{g \in \mathcal{TG}} C_g(p_g) \tag{3.10a}$$

$$s.t. \quad (3.1) \tag{3.10b}$$

$$f^p_{i'j'} = \tilde{f}^p_{i'j'} \tag{3.10c}$$

Note that the objective function only involves transmission-level resources, and the interface real power flow is fixed to the result of the forward market set-point $\tilde{f}^p_{i'j'}$, meaning that the TSO does not coordinate with the DSO.

The DSO subproblem is modeled as

$$\min_{x^D, f^p_{i'j'}} \sum_{g \in \mathcal{DG}} C_g(p_g) \tag{3.11a}$$

$$s.t. \quad (3.2) \tag{3.11b}$$

$$f^p_{i'j'} = \tilde{f}^p_{i'j'} \tag{3.11c}$$

The objective function only involves distribution-level resources, and the interface real power is fixed to the set-point of the forward market $\tilde{f}^p_{i'j'}$, implying no coordination between the DSO and the TSO.

Since the interface power flow is fixed to the result of the forward market, the operations of the TSO are fully decoupled from those of the DSO. The resulting balancing actions are feasible, provided the local imbalances do not exceed the local reserves. However, the resulting balancing action may be more costly than necessary, because the reserve resources cannot be pooled.

3.3 Numerical Illustration

This section tests the models presented in the previous sections on a small-scale system, in order to illustrate their differences. The studied network is presented in Fig. 3.2. The network consists of one transmission network and three distribution networks: the transmission network contains three buses, each of them connected to six distribution buses.

In summary[2], the system consists of two thermal units at the transmission level. The unit located in bus 1 has a marginal cost of 10 €/MWh, and a capacity of 390 MW. The unit located in bus 2 has a marginal cost of 20 €/MWh and a capacity of 150 MW. There is an inelastic demand of 350 MW in location 1. Each distribution tree has identical line characteristics and identical resources are connected to each distribution tree. Except for the roots 10, 20 and 30, each distribution bus is connected to a distributed aggregated producer of 85 MW and a distributed aggregated consumer of 80 MW. Aggregated flexible consumers with bid quantities of 50 MW and valuations ranging from 0 €/MWh up to 19.1 €/MWh are connected to each distribution bus. Thus, each of the three distribution trees can offer up to 250 MW of upward reserve (if flexible demand is fully consuming), serves a price-inelastic demand of 400 MW, and zero-cost aggregated distributed production of 425 MW (which could also offer reserve) is connected to each distribution tree.

The following discussion will concentrate on upward reserve activation, in order to keep the analysis targeted. Consider the following commitment of reserve capacity in the system, which is obtained a result of a forward reserve

[2]The full data of the model is available in the following link: http://perso.uclouvain.be/anthony.papavasiliou/public_html/Spider.dat.

capacity auction[3]: (i) Generator 2 offers 149.1 MW, at an activation cost of 20 €/MWh; (ii) Consumer 15 offers 49.1 MW, at an activation cost of 19.1 €/MWh; (iii) consumer 25 offers 6.3 MW, at an activation cost of 15 €/MWh; and (iv) consumer 34 offers 49.1 MW, at an activation cost of 19 €/MWh.

Consider an imbalance of -100 MW in bus 3, which is caused by an increase of demand in bus 3 that equals +100 MW.

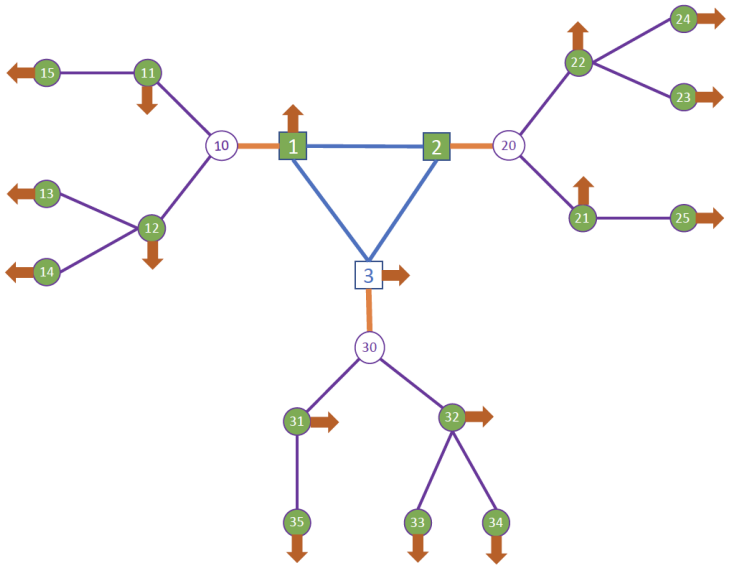


Figure 3.2: The network analyzed in the numerical illustration of section 3.3.

3.3.1 Centralized Common TSO-DSO Market

This coordination scheme resolves the imbalance by simultaneously accounting for transmission and distribution constraints. The imbalance is resolved by dispatching demand response to the following levels: (i) consumer 15 provides 49.1 MW of activated reserve[4], and its consumption level is 0.9 MW; (ii) consumer 25 offers 6.3 MW of activated reserve, its consumption level is 0.7 MW; (iii) consumer 34 offers 40.2 MW of activated reserve, its consumption level is

[3]In the present example, the reserve commitment is based on a perfectly coordinated reserve capacity auction, as described by Caramanis et al. [CNH+16].

[4]Activated reserve is measured as the difference between the production level in the co-optimization reservation problem and the production after the imbalance appears and is resolved with the co-optimization model.

9.8 MW. The real-time price becomes 19.3 €/MWh uniformly across the entire transmission network.

3.3.2 Decentralized Common TSO-DSO Market

The resolution of the DSO unperturbed model with three points of approximation of the residual supply function yields the function shown in Table 3.2. We choose a rather low number of points in order to keep the example illustrative. Note that the marginal cost function of the interface is increasing, as expected when the DSO problem is convex, and as required for market clearing of the TSO market. Note also that the residual supply function should be computed by the DSO in an open-loop fashion, without repeated rounds of information exchange, with the TSO: indeed, an iterative TSO-DSO framework is not in line with current practices or potential future practices of power system operations [GRS16].

Table 3.2: Residual supply function marginal cost (in €/MWh) for the decentralized common TSO-DSO market model.

Interface	Point 1	Point 2	Point 3
1	7.4	18.8	19.2
2	15.0	15.0	15.1
3	7.7	17.7	19.8

The resulting dispatch of the TSO resources is *identical* to the one obtained by the centralized common TSO-DSO model, with generator 1 producing 390 MW, generator 2 producing nothing, and the reserves required for balancing sourced from the interfaces. The balancing price of the transmission market is 19.7 €/MWh. It should be noted that the sourcing of the reserve shifts slightly, relative to the sourcing obtained in the centralized common market model. In the decentralized model, the interface flows balance the system with the following injections (the negative sign indicates that the flow is from the distribution pocket to the transmission system): f_1^p = -20.49 MW, f_2^p = -20.84 MW, and f_3^p = -18.67 MW. In the centralized model, the injections are f_1^p = -22.80 MW, f_2^p = -23.00 MW, and f_3^p = -14.20 MW.

One therefore observes that the decentralized approach shifts sourcing of reserve from pockets 1 and 2 to pocket 3. This can be attributed to the fact that the true marginal cost of activating distributed resources is only approximate, not exact, in the decentralized model. Notice, however, that the total amount of real power reserve which is activated should be equal in both cases, because in both cases the contribution of transmission-level resources to clearing the imbalance is identical. By consequence, there is a slight loss in consumer benefit and balancing prices at the transmission system are also slightly different from

those of the centralized common market model. Note also that, by construction, this market-clearing method will not violate system constraints.

3.3.3 Centralized Ancillary Services Market Model

This coordination scheme dispatches resources by ignoring the distribution network constraints. In particular, the imbalance is resolved by dispatching distributed demand response resources to the following levels: (i) the consumption level of consumer 15 is 13.5 MW; (ii) the consumption level of consumer 25 is 0.7 MW; and (iii) the consumption level of consumer 34 is 0.9 MW.

The real-time price becomes 19.1 €/MWh, uniformly across the entire transmission network, which is slightly lower than that of the centralized common TSO-DSO market model, and is due to the underestimation of losses. Note that there is no DLMP in this coordination scheme, instead distributed resources receive the transmission-level price.

The dispatch is not feasible, because losses in the distribution network are ignored. Solving a feasibility restoration problem, one finds that a *positive* activation of 4.0 MW is necessary in location 15 in order to restore feasibility. The results of the scheme on a toy example confirm the fact that it is not in line with the future of electricity markets. Indeed, in a configuration where distribution networks are becoming more and more important, ignoring the complex physics of the distribution network should not be considered.

Note also that a significant amount of the balancing has shifted from consumer 34 (in the centralized common TSO-DSO market model) to consumer 15 (in the centralized ancillary services model).

3.3.4 Local Ancillary Services Market Model

In this model the DSO clears a local market for reserve before a transmission-level market is cleared. The local market commits half of the reserve capacity for use by the local DSO, with the other half (the more expensive half, since the DSO reserve market clears first) being made available to the TSO. Thus, reserve capacity is allocated as follows. (i) The TSO can access reserves from generator 2 up to 149.1 MW at an activation price of 20 €/MWh, reserves from consumer 15 up to 24.55 MW at an activation price of 19.1 €/MWh, reserves from consumer 25 up to 3.15 MW at an activation price of 15 €/MWh, and reserves from consumer 34 up to 24.55 MW at an activation price of 19 €/MWh. (ii) The DSO of feeder 1 can access reserves from consumer 15 up to 24.55 MW at an activation price of 19.1 €/MWh. (iii) The DSO of feeder 2 can access reserves from consumer 25 up to 3.15 MW at an activation price of 15 €/MWh. (iv) The DSO of feeder 3 can access reserves from consumer 34 up to 24.55 MW at an activation price of 19 €/MWh.

The real-time price at the transmission level is 20.0 €/MWh. The reason is that the distribution network reserves are activated up to their full capacity (only half of the total distributed reserves are available to the TSO), and the generator with marginal cost 20 €/MWh needs to be activated. This sets the price of 20 €/MWh for balancing in the transmission network.

The imbalance is resolved by dispatching demand response resources and transmission-level generation to the following levels: (i) generator 2 produces 39.7 MW, (ii) consumer 15 withdraws 25.4 MW, (iii) consumer 25 withdraws 3.8 MW, and consumer 34 withdraws 25.4 MW. The resulting dispatch violates physical constraints, and the resolution of a phase-I feasibility restoration requires an excess production of 3.9 MW at location 34.

One notable feature of this model is that the same resource can be activated in opposite directions, depending on the imbalance for which it is activated. For example, if there is a positive transmission-level imbalance and a negative distribution-level imbalance, a distributed resource may be activated upwards by the TSO and downwards by the DSO.

As it stands, the local ancillary services market scheme demonstrates poor performances even on a toy example. This is due to the fact that we define in advance the distribution reserve capacity allocated to the transmission or the distribution network. Since this decision can be highly dependent on the market conditions, this assumption appears to be too restrictive. Based on this observation, we consider two alternatives of this scheme in chapter 5: either the TSO and the DSO activate simultaneously distribution resources (TSO has limited access to DSO resources, section 5.4.2) or the DSO solves local imbalances before giving the possibility to the TSO to activate the remaining distribution reserve capacity (Local Markets, section 5.4.3).

3.3.5 Shared Balancing Responsibility Model

This model separates the dispatch in the transmission and the distribution networks by fixing the linking variable of the two networks, which is the real power flow at the interface. The value of the real power at the interface in this numerical example is fixed to the value obtained from a forward reserve capacity auction.

The real-time price at the transmission level is 20.0 €/MWh. The reason is that the TSO has no access to distribution network reserves, and therefore the generator with marginal cost 20.0 €/MWh needs to be activated. This sets the price of 20.0 €/MWh for balancing in the transmission network. The imbalance is resolved by dispatching generator 2 at 100.0 MW.

The resulting dispatch results in a feasible power flow. This can be understood by the fact that the shared balancing responsibility model will not violate feasibility at the transmission network, because it is consistently accounting for real power (since there are no overlooked distribution network losses), and it

will not violate feasibility at the distribution network since it does not simplify the distribution network constraints. However, in general this feasibility will come at a relatively high activation cost since resources of the transmission and distribution networks are not pooled. This is evident in Table 3.3, where generator costs increase dramatically (outweighing additional consumer benefits), and exceed the activation costs of any other coordination scheme.

The relative performance of the studied coordination schemes in terms of allocative efficiency, physical feasibility, and price, are summarized in Table[5] 3.3.

Table 3.3: Summary statistics of the different coordination schemes. The coordination scheme initials stand for centralized common market (CCM), decentralized common market (DCM), centralized ancillary services market (CAS), local ancillary services market (LAS), and shared balancing responsibility (SBR).

	Gen. cost [€]	Consumer benefit [€]	Constraint violation [MW]	Transm. price [€/MWh]
CCM	3900.0	212.4	0.0	19.3
DCM	3900.0	202.2	0.0	19.7
CAS	3900.0	283.7	4.0	19.1
LAS	4693.3	1026.0	3.9	20.0
SBR	5900.0	2009.0	0.0	20.0

3.4 Conclusions

This chapter proposed models for quantifying five proposals of TSO-DSO coordination which have recently been proposed by the SmartNet consortium. The following conclusions can be drawn from the specific case study presented in this chapter: (i) The Centralized Common Market model sets the first-best standard in terms of allocative efficiency, however it is challenging to implement due to the large scale of the optimization problem and the communication requirements. (ii) The Decentralized Common Market model strikes a balance between efficiency and computational / communication tractability. This is achieved by exploiting the fact that the TSO and DSO need only agree on the amount of real power flowing over the T&D interface. (iii) The Local Ancillary Services model is dominated by the Centralized Ancillary Services model in terms of allocative efficiency. (iv) The Shared Balancing Responsibility model will not violate physical constraints, however it appears to be the least efficient solution.

[5]The degree of constraint violation is the minimum adjustment in *real* power injection / withdrawal that would be needed to restore a feasible power flow. This is why constraint violation is measured in MW.

This chapter introduced the preliminary material and thoughts on what is explored in chapters 4 and 5. We conduct a deep study and propose a framework for applying Decentralized Common Market in chapter 4. In chapter 5, we suggest two interpretations of the Local Ancillary Services model that allow for feasible dispatches, one as a non-cooperative simultaneous game and another as a Stackelberg game. We model Shared Balancing Responsibilities as a non-cooperative simultaneous game as well.

Chapter 4

Coordination of T&D System Operations in Flexibility Markets with Non-Convex Market Offers and Alternating Current Power Flows

4.1 Introduction

In this chapter, we aim at developing a practical framework for T&D balancing markets. To this end, we rely on the ideas developed in the coordination scheme called 'Decentralized Common TSO-DSO' in Chapter 3 (section 3.2.2). We extend the vision perceived in this coordination scheme and aim at validating the approach on realistic instances. Several specific features are desirable for the foreseen T&D market-clearing real-time platform and are exposed in section 4.2. We then identify the precise functioning of the market as well as the role of each actor involved (section 4.3). We present the mixed integer non-linear problem for which the platform should be able to deduce primal decisions (section 4.4) but also prices (section 4.5) in a decentralized fashion. The algorithmic details of the approaches are elaborated in section 4.6. We perform computations on realistic instances of the European markets (section 4.7) before exposing conclusions in section 4.8.

4.2 Desiderata of T&D Platforms

4.2.1 Physical Considerations

The platforms that are considered in the SmartNet test systems of Italy and Denmark include two physical complications: (i) market bids at both the transmission and distribution level can be non-convex, and (ii) the distribution network model is based on a non-convex alternating current power flow.

Non-convex market offers at the distribution level can be understood as a consequence of the disaggregated representation of resources in distribution networks. A case in point are storage resources that may be either charging or discharging. Since storage resources located in different nodes of the distribution grid need to be represented separately, the smoothing effect of aggregation vanishes. Regardless of its origin, we consider the SmartNet data as a given in our analysis, and this data included non-convex market offers at the level of the distribution system.

Linear approximations of power flow are not adequate for representing a number of relevant distribution network constraints, due to the lower voltage level at which these systems operate. Ohmic losses are especially relevant, and voltage limits can be often binding in cases of local oversupply. Reactive power flows can contribute to the stress of lines. In SmartNet [MRS+17], convex relaxations of the AC power flow equations were proposed in order to represent distribution network constraints, and linear approximations were proposed for meshed transmission networks. In the present work, we directly account for the non-convexity of the AC power flow equations in the distribution network.

4.2.2 Scalability

Integrating transmission and distribution networks naturally leads to considering market clearing problems of very large scale. Caramanis *et. al* [CNH+16] provide an indication: "*In fact, whereas transmission bus locations number in the thousands, associated distribution feeder line buses number in the hundreds of thousands or millions.*".

The SmartNet test cases additionally incorporate multi-interval look-ahead [HSZ+19]. The typical SmartNet test cases consider one-hour look-ahead, and balancing market time units of 15 minutes, hence a four-stage look-ahead where current period decisions are binding and future decisions are advisory. This market clearing is performed in a rolling horizon.

The very large scale of the problem is compounded by the challenge that a real-time / balancing dispatch and market clearing platform can only afford a few minutes of run time at best. It is therefore clear that a viable solution requires decentralization / decomposition. Compared to existing research that decomposes the problem by considering convex approximations or relaxations

[CNH+16, KCL+14], in our work we propose a decomposition of the problem that is in line with the institutional organization of power system operations, and is specifically targeted at accommodating the non-convexity of distribution network power flows. We discuss our decentralization strategy in further detail in the following section.

4.2.3 Privacy and Decentralization

The complexity and scale of the problem considered should already be sufficient reasons to consider decentralized methods for computing market clearing solutions. There are also institutional drivers for considering such decentralization. Indeed, TSOs are sometimes reluctant of undertaking the responsibility for optimizing the management of assets at the level of distribution networks, since the monitoring, optimization, control and settlement of these assets introduces overwhelming complexity to TSO operations. On the other hand, DSOs may also be reluctant of surrendering the control of assets to an entity that has no visibility of their network constraints and may therefore overload distribution network assets through the activation of flexible distribution resources.

We consider an effective communication protocol as one which allows network operators to coordinate without sharing detailed information about their local network, or the bids that are collected in their network. Such a communication protocol is compatible with various practices of bid filtering [PBDS20] or pre-qualifications [MP19], but attempts to conduct these procedures without excluding potentially high-welfare market matches. Such a protocol further exploits the typically radial structure of distribution networks, and in particular the fact that a distribution sub-network is typically connected by a single feeder to the upper-level high-voltage system. Note that, in this work, we are not exploring whether or not it is possible for the TSO or the DSO to learn information from the other's network with the framework we develop. We only assume that privacy is preserved since operators are willing to only exchange border information.

4.2.4 Consistent Pricing

The challenges discussed so far are limited to the primal problem. Since we are interested in developing a market clearing platform, we also focus on generating price signals that are consistent with the dispatch instructions generated by our platform.

Since our matching problem is non-convex, uniform market clearing prices are not guaranteed to exist. This is due both to the non-convexity of the market offers [GHP+07], as well as the non-convexity of the alternating current power flow [GNB20].

This problem is not new to electricity markets. Uplift payments are typically used for inducing agents to follow the dispatch schedule of the market. Uplift payments prevent agents from "taking matters into their own hands" by resorting to self-commitment or self-scheduling, which would limit the amount of flexible assets that are available to the network operators. In practice, the market clearing procedure is typically partitioned into two steps: (i) the primal / dispatch step matches market orders, and (ii) the dual / pricing step produces prices that are as consistent as possible to the market matches. In US markets this pass is performed once, and various methods for pricing in the second step have been considered, including IP pricing [OSH$^+$05], convex hull pricing [GHP$^+$07], approximations of convex hull pricing [HB16], and a variety of other possibilities [LA16]. In EU markets this primal/dual algortihm is performed iteratively within a branch-and-cut algorithmic scheme [MVV15], in order to accommodate additional specifications related to the management of non-convex market offers — essentially, EU DA market avoids uplifts payments by (i) constraining the problem in a way that ensures there will not be any *paradoxically accepted bids* (cleared bids actually facing losses) and (ii) allowing *paradoxically rejected bids* (a rejected bid that would have been profitable) but not paying their opportunity costs of being out of the market. Ultimately, this does not effectively reduce the uplifts to zero, but simply means increasing the total "virtual uplift" by making sure that these will not result in effective payments, due by the market operator, thanks to the market rules.

We are inspired by the decomposition of dispatch and pricing that is employed in practice. We use the amount of uplift payments, or lost opportunity cost, as a metric of how close the price signal is to being consistent with the dispatch instruction. In contrast to the aforementioned literature, the procedure that we propose aims at (i) coping with the non-convex alternating current power flow model of distribution networks, while (ii) respecting the decentralization of information that is discussed in section 4.3.2.

4.3 Overview of Market Design

The market platform that we propose aims at implementing a real-time market for clearing energy while accounting for transmission and distribution network constraints. In EU jargon, we implement an integrated real-time congestion management and balancing platform. The market dispatches flexible resources in real time so as to balance the system while ensuring that network constraints are safeguarded. The platform produces locational marginal prices for energy at the transmission level, and locational marginal prices for real and reactive power at the distribution level. Reserve capacity [CNH$^+$16] is considered as being out of scope for the present work.

Figure 4.1 provides an illustration of the sequence of electricity markets, as

well as the sequence of actions that we envision for the real-time market that we implement in our work.

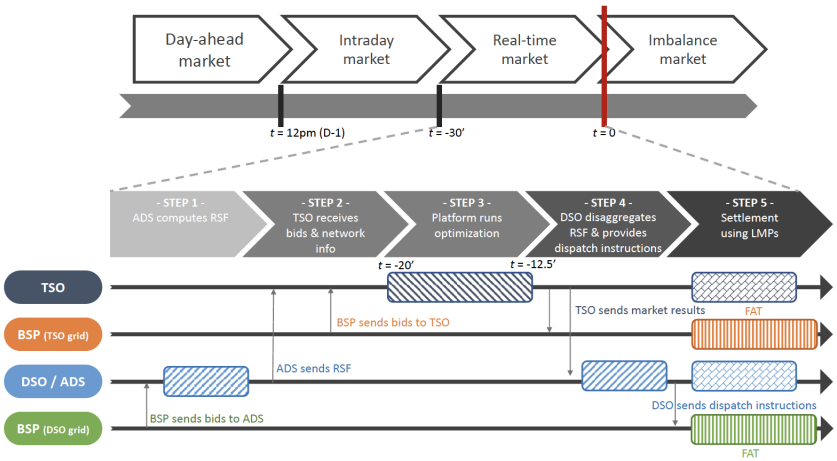


Figure 4.1: Sketch of the chronology of electricity markets with an emphasis on our vision for the sequence of events in the real-time market platform.

4.3.1 Market Agents and Products

The market platform that we develop trades location-specific real power at the transmission system, and location-specific real and reactive power at the distribution system. We assume a single meshed transmission network that is interconnected to a number of distribution networks, with each distribution network interconnected to the transmission network via a unique feeder, which we will also refer to as a transmission-distribution interface (T&D interface) (see Figure 3.1). We consider the following agents in our market model.

Transmission System Operator (TSO). The TSO procures real power from BSPs that are connected in the transmission network, as well as real power from the DSO at the transmission-distribution interface (e.g. the medium-to-high-voltage feeders). The TSO collects payments from BRPs for real power, as well as payments from the DSO for the exchange of real power at the T&D interface.

Distribution System Operator (DSO). The DSO is responsible for congestion management at the level of the distribution network. We envision a DSO that operates an active distribution system management system that is capable of producing location-specific real and reactive power prices, that are

paid to BSPs who are connected to the DSO network. The DSO collects payments for real and reactive energy from BRPs, and also pays the TSO for real power exchaged at the T&D interface.

Balancing Service Provider (BSP). BSPs operate assets that can be actively dispatched in order to resolve congestion and balance the system, i.e. they are the owners of reserve assets. This is precisely what we refer to as "flexibility". In EU jargon, we specifically refer to BSPs as owners of automatic frequency restoration reserve (aFRR) and manual frequency restoration reserve (mFRR) assets. These correspond to reserve resources that can respond within a few seconds to a few minutes. BSPs that are located in the transmission network receive location-specific balancing energy payments from the TSO. BSPs that are located in the distribution system receive payments for real or reactive balancing power from the DSO. In terms of the mathematical models that are presented in the sequel, we will identify BSPs as price-responsive "generators".

Balancing Responsible Party (BRP). Balancing responsible parties are owners of non-flexible assets, i.e. assets that are not controllable and that induce imbalances in real time. Distribution network BRPs face a location-specific price for real power and reactive power, against which their imbalances are settled. Transmission network BRPs face a location-specific price for real power. In terms of the mathematical models that are presented in the sequel, we will identify BRPs as price-inelastic "loads" at each bus of the network.

Aggregation-Disaggregation Service (ADS). The aggregation-disaggregation service implements the decentralization of transmission and distribution system operations by aggregating distribution system BSP offers, trading energy at the T&D interface, and subsequently disaggregating the interface flow to BSP activations within the distribution network (Figure 4.2). The flexibility platform also determines the location-specific real and reactive prices of the distribution network. In a centralized implementation, this flexibility platform becomes obsolete. Instead, the TSO and DSO operations are integrated into a single platform that produces dispatch signals and prices.

4.3.2 Practical Considerations

In this section, we point out certain important differences between the ideal market design described in section 4.3.1 and the existing European market design. A number of these differences can be bridged, others are rather fundamental and raise gaming aspects that extend beyond the scope of the present work.

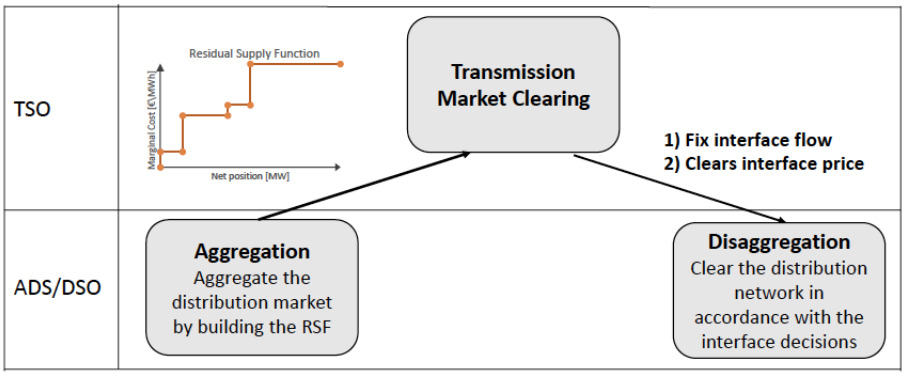


Figure 4.2: The role of the Aggregation-Disaggregation Service (ADS) in the clearing process.

In practice, TSOs are responsible for balancing and congestion management. In most EU markets, balancing energy is settled at a zonal level. The location-specific prices for real power that we propose in this work essentially implement balancing energy at nodal resolution. The EU market design faces a difficult challenge in this respect. (i) Implementing zonal flexibility markets will introduce increase-decrease (INC-DEC) gaming opportunities in these markets [HS18]. (ii) Implementing nodal real-time markets (at least in the distribution level) will introduce an inconsistency with day-ahead zonal markets that can also be exploited. These market design questions are beyond the scope of the present work.

In practice, DSOs are solely responsible for congestion management, i.e. operating the distribution system while ensuring that network constraints are respected. In our model, DSOs assume a more active role. They aggregate offers into residual supply functions that participate in the TSO-operated wholesale market. They also disaggregate feeder-level real power injections to individual BSP assets that are connected in the distribution network.

In European markets, BRPs are typically exposed to an imbalance price that is not location-specific. Instead, in our analysis we assume that BRP positions are settled at locational prices, and that BRP net positions are measurable at nodal resolution (if not in real time, at least ex post for the purpose of settlement).

One final relevant remark is that imbalance is typically monitored at the level of a load frequency control (LFC) area. Thus, imbalances cannot be disaggregated, in real time, at the level of distribution nodes. Thus, BRPs can only be charged after the fact for the imbalances that they induce on the network. In our model we assume imbalances that are measurable at nodal resolution in real time.

4.3.3 Residual Supply Functions

In order to bridge the gap between TSO-DSO coordination and decentralized operations of the network operators, we propose in this chapter an Aggregation-Disaggregation Service (ADS) which relies on Residual Supply Functions (RSFs) [PM18, PBDS20]. In the convex case, RSFs correspond to a Benders decomposition of the TSO-DSO coordination problem, and are thus distinct from the dual decomposition methods set forth in [KCL+14, CNH+16]. We opt for RSFs as being better aligned with the institutional operation of real-time EU markets, as we discuss in the following.

Concretely, one RSF is defined for every T&D interface. It corresponds to the gradient of the least cost at which a given amount of real power can be exported from a distribution network to the T&D interface. An important advantage of RSFs is their direct interpretation as bids in the balancing market operated by the TSO, which mask the complexity of the underlying distribution network. The ADS computes an RSF by collecting the market bids of all BSPs that are connected to the distribution network of a given DSO, as well as the relevant distribution network parameters from the DSO. The RSF can then be bid into the wholesale market, and compete on equal footing with the BSP offers of transmission-level BSPs. The clearing of the market implies a net position for the DSO. This net position needs to be disaggregated to individual BSPs at the distribution network by the disaggregation function of the ADS.

Note that this idea can be generalized to any hierarchical structure between operators and for example coordination of national TSOs in the European context [PBDS20].

Example of the use of the RSF by the ADS. The best way to understand the functioning of the RSF is with an example. Let us consider the example in Figure 4.3 accompanied with Table 4.1 displaying generation capacities and costs.

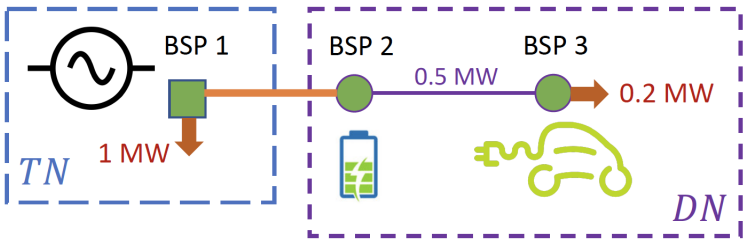


Figure 4.3: Simple T&D example.

Note that we assume neither losses nor reactive power in the distribution network for simplicity and that there is a 0.5 MW real power flow limit on the

Table 4.1: Data of T&D example in Figure 4.3.

BSP	Capacity	Marginal Cost
BSP 1	3 MW	20 €/MWh
BSP 2	1 MW	15 €/MWh
BSP 3	1 MW	10 €/MWh

line connecting BSP 2 and BSP 3. The imbalances to cover are a 1 MW load located on the single transmission bus and a 0.2 MW load at bus 3. These imbalances can either be covered by the expensive transmission BSP (BSP 1) or by the cheaper distributed renewable resources (BSP 2 and BSP 3).

In a centralized setting, the solution would be the following:

- BSP 1: activate 0 MW, clearing price 15€/MWh.
- BSP 2: activate 0.5 MW, clearing price 15€/MWh.
- BSP 3: activate 0.7 MW, clearing price 10€/MWh.

From the practical considerations reported in section 4.3.2, it is necessary to consider a decentralized solution framework. The first task of the ADS is to aggregate the distribution flexibility. This is easily done by considering the distribution network as only one bus at the transmission level. This is represented in Figure 4.4.

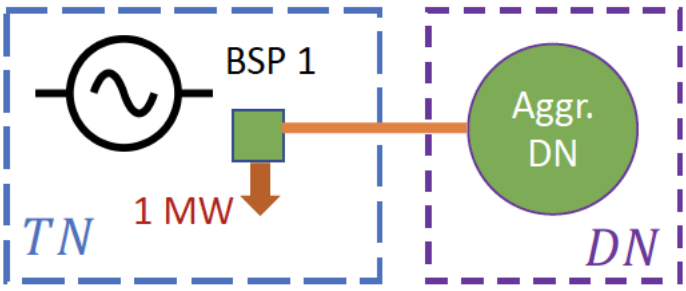
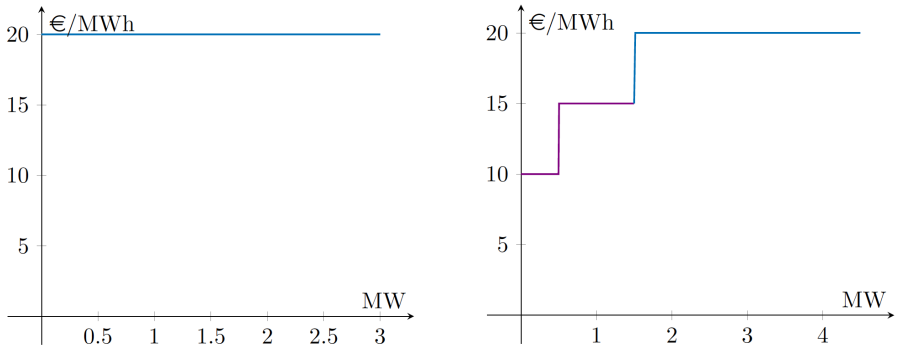


Figure 4.4: Aggregation from the ADS of the example in Figure 4.3.

The aggregation of the distribution network would need to capture the BRP imbalances, BSP offers and physics of the distribution networks while sharing a minimal amount of information with the TSO. The goal of the work is to show how the RSF can guarantee this property. If the TSO does not take into account the distribution network, the only possibility is to activate BSP 1 (Figure 4.5(a)). The role of the ADS is then to build the RSF by assuming different levels of interface flows. In this example, considering for instance 0

MW, 0.5 MW, 1 MW, 1.5 MW and 2 MW interface flow levels provides an exact representation of the RSF. What the TSO is facing for covering the imbalance is represented in Figure 4.5(b): the violet part of the curve is the RSF while the blue part of the curve represents BSP 1.



(a) TSO with only transmission information.

(b) TSO with ADS sharing the RSF.

Figure 4.5: On the example of the Figure 4.3, depending on the willingness to collaborate, the TSO faces the following options in (a) the 'only transmission information' case and (b) the case where the ADS shares the RSF.

In the transmission balancing market, the system imbalance is matched against +1 MW of the RSF. The TSO is then buying 1 MW to the ADS at the price indicated by the RSF of 15€/MWh. The primal/dispatch part of the disaggregation function then runs, with an aim of finding the most efficient way to evacuate 1 MW while respecting network constraints and distribution imbalances (−0.2 MW at bus 3). The cheapest way for this to be achieved is by activating BSP2 and BSP3 upward by 0.5 MW and 0.7 MW respectively. The dual/pricing part of the disaggregation function then runs, with an aim at determining prices that are coherent with the dispatch and the balancing price. Concretely, we arrive at a price of 15 €/MWh at bus 2 and 10 €/MWh at bus 3. The interface price of 15 €/MWh is consistent with the partial activation of BSP2, which is indifferent about any quantity of activation, since it is making a zero-profit margin. This consistency pricing step is discussed in more details in section 4.5.2 and matches Step 6 of Algorithm 6.

Aiming at precisely representing the financial interactions between the actors in a decentralized market clearing platform, the resulting settlement table of the example is drawn in Table 4.2.

In Table 4.2, we report the financial position of the different transmission and distribution actors, presented in the columns of the table. The settlement is divided in 4 phases:

Table 4.2: Settlement table of the example of Figure 4.3.

	Transmission			Distribution			
Settlements	BSP 1	BRP 1	TSO	ADS	BSP 2	BSP 3	BRP 3
TM-BSP	0 €	-	−15 €	+15 €	-	-	-
Quantity (MW)			−1	+1			
Price (€/MWh)			15	15			
ADS Dis.	-	-	-	−14.5 €	+7.5 €	+7 €	-
Quantity (MW)				−1.2	+0.5	+0.7	
Price (€/MWh)				[10;15]	15	10	
TM-BRP	-	−15 €	+17 €	-	-	-	−2 €
Quantity (MW)		−1	+1.2				−0.2
Price (€/MWh)		15	[10;15]				10
ADS Rebal.	-	-	−2 €	+2 €	-	-	-
Quantity (MW)			−0.2	+0.2			
Price (€/MWh)			10	10			
Total	0 €	−15 €	0 €	+2.5 €	+7.5 €	+7 €	−2 €

1. Transmission market BSP (TM-BSP) settlement. This is the transmission market cleared by the TSO when taking into account transmission BSP offers and the virtual distribution bid sent by the ADS, i.e. the RSF.

2. ADS disaggragation (ADS Dis.). The ADS determines the distribution BSP activations in accordance with the interface power flow cleared in the TM-BSP.

3. Transmission market BRP (TM-BRP) settlement. The TSO collects transmission and distribution BRP payments with respect to the LMPs cleared in the previous steps.

4. ADS Rebalancing (ADS Rebal.). The TSO reallocates the distribution BRP payments received to compensate the ADS which is in charge of managing the congestion in the distribution network with the DSO.

In the first phase, the TM-BSP, the TSO requests 1 MW to the ADS which is bought at a price of 15€/MWh without activating transmission BSP offers. The ADS then disaggregates the interface flow by requesting the activation of BSP 2 and BSP 3 who are remunerated according to the locational marginal prices (respectively 15€/MWh and 10€/MWh). Note that, in the same time, the ADS covers for distribution BRP imbalance: here a 0.2 MW-load located at bus 3, which pushes the ADS to request a total of 0.7 MW from BSP 3. The next step consists in the BRP settlement in which the TSO collects payments from transmission and distribution BRPs (BRP 1 and BRP 3) who are facing locational prices (15€/MWh at bus 1, 15€/MWh at bus 3). Finally, the TSO redistributes the payment received by the distributed BRPs (here BRP 3) to

the ADS. The last row shows what each participant paid or got paid. As expected, BSPs and BRPs are remunerated according to their consumption and the locational marginal price cleared. The TSO and the ADS shares the congestion rent which appears through price differences at each location. The congestion rent is here fairly assigned: the TSO is not facing any price difference in its network while the ADS has a congested line (the line connecting buses 2 and 3) and a price difference between these two buses. The ADS receives a compensation of $+2.5€$ ($= (15 - 10) \times 0.5$).

In practice, computing an exact representation of the RSF is not tractable because of the complexity of the distribution networks. We show in this work how approximations of the RSF provide satisfactory results. We also point out that the RSF is only relevant for the decentralized version of our proposed market clearing platform, which is described in section 4.6.4. The centralized versions of the market clearing platform that are described in sections 4.6.1, 4.6.2 and 4.6.3 do not rely on the RSF.

4.4 The Primal Market Clearing Problem

In this section we gradually build up the market clearing problem that we aim at solving. We specifically focus on providing an accurate account of the non-convexities of the primal problem.

4.4.1 Multi-Period Optimal Power Flow

We now extend the OPF_{AC} introduced in Chapter 1 (problem (1.5)) on a time horizon $\mathcal{T} = \{1, \dots, t_f\}$. In this setting, power balance constraints must be satisfied at each time step. The set of operational constraints, $\mathcal{OC}$, is also valid at each time step and we assume that the parameters of the network do not change from one time-step to another. Consequently, operational constraints need to be met at each time-step and do not create time-coupling. On the contrary, the set of generation constraints, $\mathcal{GC}_g$, is time-dependent (for example generation capacity might change) and involves time-coupling. Indeed, when considering multi-period OPF on a time horizon $\mathcal{T} = \{1, \dots, t_f\}$, $\mathcal{GC}_g$ could also include inter-temporal constraints, such as ramp constraints (among others):

$$\underline{RP}_{g,t} \leq p_{g,t+1} - p_{g,t} \leq \overline{RP}_{g,t}, \ \forall g \in \mathcal{G}, t \in \mathcal{T} - \{t_f\}.$$

In the multi-period context, we keep the same notation $\mathcal{GC}_g$ for generator $g \in \mathcal{G}$.

The multi-period OPF on a time-horizon $\mathcal{T} = \{1, \dots, t_f\}$ is then formulated

as follows:

$$\min_{x} \quad \sum_{g\in\mathcal{G}}\sum_{t\in\mathcal{T}} C_{g,t}(p_{g,t}) \tag{4.1a}$$

$$s.t. \quad F_{i,t}(x_t) = 0, \qquad \forall i \in \mathcal{B},\ \forall t \in \mathcal{T} \tag{4.1b}$$

$$(v, \theta, f_t^p, f_t^q) \in \mathcal{OC}, \qquad \forall t \in \mathcal{T} \tag{4.1c}$$

$$(p_g, q_g) \in \mathcal{GC}_g, \qquad \forall g \in \mathcal{G} \tag{4.1d}$$

In (4.1), the coupling constraints appear in (4.1d). Note that the multi-period OPF can also be extended to the other formulations (quadratic AC-OPF, quadratic relaxation, SOCP relaxation and DC approximation). The problem remains a continuous non-convex optimization problem.

4.4.2 Non-Convex Offers

The bids that we consider in this chapter are based on the market product specifications of the SmartNet EU project. They thus obey the structure of the bids used in the single day-ahead European market coupling model. The detailed description of the bid structure can be found in [LPG+19](Section 3).

To incorporate complex bids (like block bids, linked bids etc.), we introduce $d_g \geq 1$ binary variables y_g for each generator g. The set $\mathcal{GC}_g$ is now extended and can incorporate binary constraints:

$$(p_g, q_g, y_g) \in \mathcal{GC}_g, \qquad \forall g \in \mathcal{G} \tag{4.2a}$$

$$y_g \in \{0,1\}^{d_g}, \qquad \forall g \in \mathcal{G} \tag{4.2b}$$

The set $\mathcal{GC}_g$ can incorporate various features. These include linked bids, exclusive acceptance, minimum acceptance duration, and so on. These bids therefore introduce binary variables and inter-temporal constraints. For more details concerning the bid structure and a precise mathematical formulation, the reader is referred to appendix 4.A.

4.4.3 The Integrated T&D Real-Time Market Clearing Primal Model

The topology of the networks considered in this chapter is presented in Figure 3.1. We display a single distribution network in the figure, whereas the test systems in our case study include multiple distribution networks that are connected to transmission network buses.

We denote by $\mathcal{B} = \mathcal{TB} \cup \mathcal{DB}$ the set of buses which is the union of the set of transmission buses and distribution buses. Similarly, the set of generators $\mathcal{G} = \mathcal{TG} \cup \mathcal{DG}$ is the union of the set of transmission generators and the set of distribution generators, and $\mathcal{L} = \mathcal{TL} \cup \{l' = (i', j')\} \cup \mathcal{DL}$ is the set of

lines, which is the union of the transmission lines, interconnection line and distribution lines.

The transmission network is meshed. We employ a DC approximation for this network [SA74], following the SmartNet market clearing model. Non-binding voltage constraints and lower importance of real power losses are common assumptions when considering transmission power systems [FSR12,PM18]. Like in Chapter 3, we assume that the interconnection is a lossless line and only model the interconnection real power flow by using one variable $f^p_{l'}$ and duplicate it in both transmission ($f^{p,T}_{l'}$) and distribution networks ($f^{p,D}_{l'}$). The distribution network is assumed to be radial where the SOC relaxation is often exact in practice. Nevertheless, physically feasible dispatch decisions are considered a requirement for our market clearing platform, therefore we will require that the AC power flow equations are satisfied in the distribution network. Since the network is radial, we will make the use of the quadratic relaxation, OPF_Q, proven to be exact under radial topology assumption [ZT11]. Similar requirements can be imposed for the transmission network, but are out of scope for the present work. Note that such AC feasibility requirements for the transmission network are not imposed on the SmartNet instances that are solved in the case study (and corresponding data is not available for the test systems of these case studies).

The formulation of the integrated T&D market clearing problem is as follows:

$$\min_{(x,y,f^p_{l'})} \quad \sum_{g\in\mathcal{G}}\sum_{t\in\mathcal{T}} C_{g,t}(p_{g,t}) \tag{4.3a}$$

$$s.t. \quad F^{DC}_{i,t}(x^T_t) = 0, \quad (\lambda_{i,t}) \quad \forall i \in \mathcal{TB}, \forall t \in \mathcal{T} \tag{4.3b}$$

$$(\theta_t, f^p_{\mathcal{TL},t}) \in \mathcal{OC}^{DC}, \quad \forall t \in \mathcal{T} \tag{4.3c}$$

$$(p_g, y_g) \in \mathcal{GC}^{DC}_g, \quad \forall g \in \mathcal{TG} \tag{4.3d}$$

$$f^p_{l',t} = f^{p,T}_{l',t}, \quad \forall t \in \mathcal{T} \tag{4.3e}$$

$$f^p_{l',t} = f^{p,D}_{l',t}, \quad \forall t \in \mathcal{T} \tag{4.3f}$$

$$f^p_{l',t} \in \mathcal{IOC}, \quad \forall t \in \mathcal{T} \tag{4.3g}$$

$$F^Q_{i,t}(x^D_t) = 0, \quad (\lambda_{i,t}) \quad \forall i \in \mathcal{DB}, \forall t \in \mathcal{T} \tag{4.3h}$$

$$(c_t, s_t, f^p_{\mathcal{DL},t}, f^q_t) \in \mathcal{OC}^Q, \quad \forall t \in \mathcal{T} \tag{4.3i}$$

$$(p_g, q_g, y_g) \in \mathcal{GC}_g, \quad \forall g \in \mathcal{DG} \tag{4.3j}$$

$$y_g \in \{0,1\}^{d_g}, \quad \forall g \in \mathcal{G} \tag{4.3k}$$

$$x^T = (p_{\mathcal{TG}}, f^p_{\mathcal{TL}}, \theta, f^{p,T}_{l'}),\ x^D = (p_{\mathcal{DG}}, q, c, s, f^p_{\mathcal{TL}}, f^q, f^{p,D}_{l'})$$

This problem is a mixed integer non-linear problem and its continuous relaxation is non-convex. Integrality comes from the introduction of binary vari-

ables y. In addition to binary variables, non-convexities originate from constraint (4.3i) modeling power flow equations in the distribution network. Intertemporality appears in two constraints: (4.3d) and (4.3j). Note that, even if we do not explicitly define them in problem (4.3), slack variables are introduced and highly penalized to ensure feasibility of the problem. These slack variables can be seen as production or load shedding and are common when modeling OPF problems. These slack variables are also present in the problems we introduce in section 4.6 which ensure feasibility of the of the approaches that we develop. Solution methods for solving (4.3) are presented in section 4.6.

4.5 Pricing

The model that we present in section 4.4.3 constitutes the primal dispatch problem. Our goal is to develop a platform that, in addition to dispatch decisions, also computes market clearing prices that can support a competitive equilibrium. Since the primal problem is non-convex (in both the network model as well as the market orders), there is no guarantee that such uniform prices exist. The approaches we developed largely rely on O'Neill pricing [OSH$^+$05] and are detailed in section 4.6. The general idea of O'Neill pricing is to (i) solve a certain problem having binary and continuous variables; (ii) fix the binary variables to their optimal value; (iii) solve the resulting continuous problem and deduce prices as dual variables of the constraints of interest. Here the variables of interest are transmission and distribution power balance constraints (4.3b)–(4.3h) and prices are denoted λ (λ^p (resp. λ^q) associated with active (resp. reactive) locational marginal price). The main objective of this work is to provide a decentralized real-time market clearing framework capable of overcoming the challenges mentioned in section 4.2. A discussion on finding out which pricing scheme should be chosen is out of the scope of this work but such discussion is definitely of importance in the T&D coordination and market-clearing analysis.

As a metric of the quality of our derived prices, we introduce lost opportunity costs [SZZL15, HSZ$^+$19], which we define precisely in the following.

4.5.1 Lost Opportunity Cost

Computing the lost opportunity cost (LOC) is a way of ensuring that the primal solution (dispatch variables) and the dual solution (prices) are consistent and fair in the sense of aligning agents' selfish profit-maximizing actions with the market dispatch. We differentiate two types of LOC: one for the generators, and one for the network. Generators' LOC would be called *generator side-payment* in [GNB20] and the network LOC is called *potential congestion revenue shortfall*. We detail the exact optimization problem that is solved for computing the LOC in both cases. We assume that the market platform has

produced dispatch solutions $\hat{x}$, binary solutions $\hat{y}$, and market prices $\hat{\lambda}$.

4.5.1.1 Generator / BSP

When facing the electricity price $\hat{\lambda}$, the goal of the generator is to maximize profit. Concretely, assuming that $\hat{x}$, $\hat{y}$ and $\hat{\lambda}$ have been computed, the *potential* profit of a generator $g \in \mathcal{G}$ can be computed by solving the following optimization problem:

$$\max_{(p_g, q_g, y_g)} \quad \text{Profit}_g(p_g, q_g, y_g; \hat{\lambda}) = \sum_{t \in \mathcal{T}} \left(\hat{\lambda}^p_{\mathcal{B}(g),t} p_{g,t} - C_{g,t}(p_{g,t}) + \hat{\lambda}^q_{\mathcal{B}(g),t} q_{g,t} \right) \tag{4.4a}$$

$$s.t. \qquad (p_g, q_g, y_g) \in \mathcal{GC}_g \tag{4.4b}$$

$$y_g \in \{0,1\}^{d_g} \tag{4.4c}$$

Note that, as we indicate in section 4.3.1, the generator faces a locational marginal price for real and reactive power.

Here, $\mathcal{B}(g)$ is the bus where generator g is located. Denote the optimal solution of problem (4.4) as $(\tilde{p}_g, \tilde{q}_g, \tilde{y}_g)$. Then the selfish profit of generator g when facing price $\hat{\lambda}$ is $\text{Profit}_g(\tilde{p}_g, \tilde{q}_g, \tilde{y}_g; \hat{\lambda})$. This profit has to be compared to the profit induced by the dispatch decisions sent by the market operator: $\text{Profit}_g(\hat{p}_g, \hat{q}_g, \hat{y}_g; \hat{\lambda})$. The lost opportunity cost of generator g is then defined as follows:

$$\text{LOC}_g(\hat{p}_g, \hat{q}_g, \hat{y}_g, \hat{\lambda}) = \text{Profit}_g(\tilde{p}_g, \tilde{q}_g, \tilde{y}_g; \hat{\lambda}) - \text{Profit}_g(\hat{p}_g, \hat{q}_g, \hat{y}_g; \hat{\lambda})$$

This value can be interpreted as how much the generator could have gained by deviating from the dispatch instructions sent by the market operator. This constitutes an opportunity cost for the generator, and is a measure of how "inconsistent" the prices of the market clearing platform are, relative to the dispatch instructions. Problem (4.4) can easily be adapted in case the generator is connected to the transmission network and is not concerned with reactive power injections.

4.5.1.2 Network Operator

The network operator decision variables are (c, s, f^p, f^q) when considering AC lines (extending this concept to a T&D network is straightforward). In the same spirit as for the generator, we can measure the lost opportunity cost of the network operator by considering the potential arbitrage profit that the network operator can achieve by trading real and reactive power in different

parts of the network:

$$\max_{(c,s,f^p,f^q)} \quad \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{B}}\left(\hat{\lambda}^p_{i,t}(-G_i c_{ii,t} - \sum_{j\in\delta(i)} f^p_{ij,t}) + \hat{\lambda}^q_{i,t}(B_i c_{ii,t} - \sum_{j\in\delta(i)} f^q_{ij,t})\right) \tag{4.5a}$$

$$s.t. \qquad (c_t, s_t, f^p_t, f^q_t) \in \mathcal{OC}^Q, \qquad \forall t \in \mathcal{T} \tag{4.5b}$$

Since this problem is non-linear and non-convex as well as large scale, in our numerical experiments we will compute the *relaxed* LOC which is based on the SOC relaxation of the OPF. This allows us to upper bound the LOC in our numerical experiments, which is relevant for our goal of assessing the quality of the solution produced by our market clearing platform. The relaxed profit maximization of the network operator can be expressed as follows:

$$\max_{(c,s,f^p,f^q)} \quad \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{B}}\left(\hat{\lambda}^p_{i,t}(-G_i c_{ii} - \sum_{j\in\delta(i)} f^p_{ij,t}) + \hat{\lambda}^q_{i,t}(B_i c_{ii} - \sum_{j\in\delta(i)} f^q_{ij,t})\right) \tag{4.6a}$$

$$s.t. \qquad (c_t, s_t, f^p_t, f^q_t) \in \mathcal{OC}^{SOC}, \qquad \forall t \in \mathcal{T} \tag{4.6b}$$

The solution of problem (4.6) is denoted as $(\tilde{c}, \tilde{s}, \tilde{f}^p, \tilde{f}^q)$ and the objective value of the problem as $\text{Profit}_{\text{NO}}(\tilde{c}, \tilde{s}, \tilde{f}^p, \tilde{f}^q; \hat{\lambda})$. The LOC for the network operator is then defined as follows:

$$\text{LOC}_{\text{NO}}(\hat{c}, \hat{s}, \hat{f}^p, \hat{f}^q, \hat{\lambda}) = \text{Profit}_{\text{NO}}(\tilde{c}, \tilde{s}, \tilde{f}^p, \tilde{f}^q; \hat{\lambda}) - \text{Profit}_{\text{NO}}(\hat{c}, \hat{s}, \hat{f}^p, \hat{f}^q; \hat{\lambda})$$

While LOC_g reflected the consistency of the price at each node with respect to the dispatch instructions of the generators, LOC_{NO} reflects the consistency of the market prices between the nodes themselves: the network operator is facing the market prices and is willing to maximize the value of the network with respect to these prices. In his perspective, each"bilateral trad" between two nodes will generate a congestion revenue but will also occupy spaces on the grid, generating a lost opportunity cost linked to the other trades that it prevents. The prices and dispatch would form an equilibrium with respect to LOC_{NO} if, given the prices, there is no opportunity of arbitrage left that could increase the revenue out of the network.

In actual markets, this shortfall corresponds to potential financial exposure of the network operator from selling financial transmission rights [GNB20].

The total market LOC can be computed as follows, and compared to the depth of the market, which we quantity using the sum of generator revenues and load payments that we call producer load payments (PLP). Concretely,

given a market clearing solution $(\hat{x}, \hat{y}, \hat{\lambda})$, we define the following:

$$\mathrm{LOC}(\hat{x}, \hat{y}, \hat{\lambda}) = \mathrm{LOC}_{\mathrm{NO}}(\hat{c}, \hat{s}, \hat{f}^p, \hat{f}^q, \hat{\lambda}) + \sum_{g \in \mathcal{G}} \mathrm{LOC}_g(\hat{p}_g, \hat{q}_g, \hat{y}_g, \hat{\lambda})$$

$$\begin{aligned}\mathrm{PLP}(\hat{x}, \hat{y}, \hat{\lambda}) = &\sum_{g \in \mathcal{G}} \sum_{t \in \mathcal{T}} \max(\hat{p}_{g,t}, 0) \hat{\lambda}^p_{\mathcal{B}(g),t} + \sum_{i \in \mathcal{B}} \sum_{t \in \mathcal{T}} \max(-D^p_{i,t}, 0) \hat{\lambda}^p_{i,t} \\ &+ \sum_{g \in \mathcal{DG}} \sum_{t \in \mathcal{T}} \max(\hat{q}_{g,t}, 0) \hat{\lambda}^q_{\mathcal{B}(g),t} + \sum_{i \in \mathcal{DB}} \sum_{t \in \mathcal{T}} \max(-D^q_{i,t}, 0) \hat{\lambda}^q_{i,t}\end{aligned}$$

In the results section, we report the values LOC and PLP (section 4.7.2, table 4.5). The LOC should be reported relative to the PLP.

4.5.2 Decentralized Computation of Dual Optimal Multipliers

As we explain in section 4.3.2, our interest is in preserving a decentralized framework. This applies both for the computation of the primal solution, as well as for the computation of market clearing prices. One aspect of pricing in a decentralized fashion is ensuring price consistency, in our case between the transmission and the distribution network. We comment on the decentralized computation of prices (or dual optimal multipliers) using the example of section 4.3.3 and Figure 4.3.

Example. Coming back to the example in Figure 4.3, let us first write a simplified version of the centralized problem (DC OPF approximation in the distribution network and we do not model angles for simplicity):

$$\begin{array}{lll}
\min & 20p_1 + 15p_2 + 10p_3 & \\
s.t. & p_1 - 1 = f^p_{12} & (\lambda_1) \\
& p_2 = -f^p_{12} + f^p_{23} & (\lambda_2) \\
& p_3 - 0.2 = -f^p_{23} & (\lambda_3) \\
& 0 \leq p_1 \leq 3 & (\eta^-_1, \eta^+_1) \\
& 0 \leq p_2 \leq 1 & (\eta^-_2, \eta^+_2) \\
& 0 \leq p_3 \leq 1 & (\eta^-_3, \eta^+_3) \\
& -0.5 \leq f^p_{23} \leq 0.5 & (\gamma^-_{23}, \gamma^+_{23})
\end{array}$$

The primal optimal decision $p_1 = 0,\ p_2 = 0.5,\ p_3 = 0.7,\ f_{12}^p = -1,\ f_{23}^p = -0.5$ is easily obtained and the KKTs linking the dual variables are the following:

$$\begin{aligned}
&\eta_1^+ = \eta_2^- = \eta_2^+ = \eta_3^- = \eta_3^+ = \gamma_{23}^+ = 0\\
&-20 + \lambda_1 + \eta_1^- = 0\\
&-15 + \lambda_2 = 0\\
&-10 + \lambda_3 = 0\\
&-\lambda_1 + \lambda_2 = 0\\
&-\lambda_2 + \lambda_3 + \gamma_{23}^- = 0
\end{aligned}$$

We then have 11 dual variables defined in a system of 11 linear equations which leads to the following unique dual solution for this problem:

$$\begin{aligned}
&\lambda_1 = 15,\ \lambda_2 = 15, \lambda_3 = 10\\
&\eta_1^- = 5,\ \eta_1^+ = \eta_2^- = \eta_2^+ = \eta_3^- = \eta_3^+ = 0\\
&\gamma_{23}^- = 5,\ \gamma_{23}^+ = 0
\end{aligned}$$

The λ vector provides the LMPs obtained in a centralized fashion and we also need to compute it in a decentralized fashion.

In section 4.3.3, we showed how it is possible to obtain the primal optimal solution if the RSF is correctly built. In particular, assuming that the interface real power flow decision, f_{12}^p, is previously fixed to $f_{12}^p = -1$, the transmission problem to solve in order to obtain the transmission clearing decisions is the following:

$$\begin{aligned}
&\min && 20p_1 && && (4.7a)\\
&s.t. && p_1 - 1 = f_{12}^p, && (\lambda_1) && (4.7b)\\
&&& 0 \leq p_1 \leq 3, && (\eta_1^-, \eta_1^+) && (4.7c)\\
&&& f_{12}^p = -1 && (\mu_{12}) && (4.7d)
\end{aligned}$$

The primal solution of this problem is $\hat{p}_1 = 0, \hat{f}_{12}^p = -1$ and is in line with the centralized solution. Consequently, we have the following KKT conditions to deduce the dual variables:

$$\begin{aligned}
&\eta_1^+ = 0\\
&-20 + \lambda_1 + \eta_1^- = 0,\ \eta_1^- \geq 0\\
&-\lambda_1 + \mu_{12} = 0
\end{aligned}$$

This system of equations has 4 variables and 3 equality constraints which results in a non-unique dual solution. For example, solving problem (4.7) with a commercial solver, such as Mosek, leads to $\lambda_1' = 20, \mu_{12}' = 20, \eta_1^{-\prime} = 0, \eta_1^{+\prime} = 0$.

In particular, the price cleared at bus 1 ($\lambda'_1 = 20$) is different from the centralized price ($\lambda_1 = 15$) due to multiple dual optimal solutions of problem (4.7).

To overcome this potential issue, we suggest to solve a supplementary transmission problem to compute the prices. We use an additional information on the RSF: the price associated to the flow level $\hat{f}^p_{12} = -1$. From Figure 4.5, this flow level has a marginal cost of $\hat{\mu}_{12} = 15$€/MWh for the transmission network. To recover the LMP λ_1, we solve the following problem:

$$\begin{aligned}
&\min && 20p_1 - 15f^p_{12} && && (4.8a)\\
&s.t. && p_1 - 1 = f^p_{12}, && (\lambda_1) && (4.8b)\\
& && 0 \leq p_1 \leq 3, && (\eta^-_1, \eta^+_1) && (4.8c)
\end{aligned}$$

This problem has the same primal solution (even if it is not the case in general) and the KKTs are then:

$$\begin{aligned}
&\eta^+_1 = 0\\
&-20 + \lambda_1 + \eta^-_1 = 0,\ \eta^-_1 \geq 0\\
&-\lambda_1 + 15 = 0
\end{aligned}$$

The same LMPs as the ones obtained centrally are obtained in a decentralized fashion. The same idea can be applied to the distribution problem to obtain the distribution dispatch decisions and LMPs.

To highlight how crucial it is to recover consistent prices, we consider the pricing solution we obtained from directly solving (4.7) and the one we obtained from solving additionally (4.8) for the same primal optimal solution, $(\hat{p}, \hat{f}^p)$. We assume that the distribution LMPs are the same ($\hat{\lambda}_2 = \lambda'_2 = 15$ and $\hat{\lambda}_3 = \lambda'_3 = 10$), the only difference comes from the transmission price: $\hat{\lambda}_1 = 15$ and $\lambda'_1 = 20$. If one were to compute the generator LOC, LOC_g for each BSP g, one would see that both pricing vectors $\hat{\lambda}$ and λ' incentize BSPs to participate in the market because it leads to a zero lost opportunity cost. Let us compute the network operator LOC, LOC_{NO}, for both LMPs $\hat{\lambda}$ and λ' (we contract $\text{Profit}_{\text{NO}}$ to P_{NO} here):

$$\begin{aligned}
\text{LOC}_{\text{NO}}(\hat{f}^p_{12}, \hat{\lambda}) &= \max_{\substack{f^p_{12}\\ -0.5 \leq f^p_{23} \leq 0.5}} \left(-\hat{\lambda}_1 f^p_{12} + \hat{\lambda}_2(f^p_{12} - f^p_{23}) + \hat{\lambda}_3 f^p_{23}\right) - \text{P}_{\text{NO}}(\hat{f}^p_{12}, \hat{\lambda})\\
&= \max_{\substack{f^p_{12}\\ -0.5 \leq f^p_{23} \leq 0.5}} (-15f^p_{12} + 15(f^p_{12} - f^p_{23}) + 10f^p_{23}) + 5\hat{f}^p_{23}\\
&= \left(\max_{-0.5 \leq f^p_{23} \leq 0.5} -5f^p_{23}\right) - 2.5 = 0
\end{aligned}$$

$$\begin{aligned}
\mathrm{LOC}_{\mathrm{NO}}(\hat{f}^p_{12}, \lambda') &= \max_{\substack{f^p_{12} \\ -0.5 \le f^p_{23} \le 0.5}} \quad (-\lambda'_1 f^p_{12} + \lambda'_2 (f^p_{12} - f^p_{23}) + \lambda'_3 f^p_{23}) - \mathrm{P}_{\mathrm{NO}}(\hat{f}^p_{12}, \lambda') \\
&= \max_{\substack{f^p_{12} \\ -0.5 \le f^p_{23} \le 0.5}} \quad (-20 f^p_{12} + 15(f^p_{12} - f^p_{23}) + 10 f^p_{23}) + 5\hat{f}^p_{12} + 5\hat{f}^p_{23} \\
&= \left(\max_{\substack{f^p_{12} \\ -0.5 \le f^p_{23} \le 0.5}} \quad -5 f^p_{12} - 5 f^p_{23} \right) - 7.5 = +\infty
\end{aligned}$$

As expected, since the prices obtained by solving additionally (4.8) match the ones computed centrally, $\mathrm{LOC}_{\mathrm{NO}}(\hat{f}^p_{12}, \hat{\lambda})$ is 0. But, if one does not recover consistent prices, it leads to an infinite LOC for the network operator. A high LOC is the consequence of a wrong valuation of the electricity price at certain locations. Consistent prices are necessary to signal coherent investment decisions. Indeed, investments would typically target locations for which the price of electricity is high in certain configurations. If the prices are wrongly addressed, it may lead to undesirable signals.

These ideas form the basis of our reasoning in section 4.6.4 and Algorithm 6.

4.6 Proposed Market Clearing Algorithms

We present three centralized configurations and compare them with one decentralized configuration for producing market clearing matches and prices. We motivate each of the approaches, and subsequently discuss in section 4.7 the results that they produce on realistic instances of the integrated T&D market clearing problem.

4.6.1 Approach 1: Relaxation

The idea of Approach 1 is to implement the most straightforward centralized scheme so as to derive a relaxed solution of the problem as well as a lower bound (LB) of the optimal solution. In order to achieve this requirement, we first solve the complete problem by relaxing the AC power flow equations in the distribution network and make use of the SOC relaxation. The problem to

solve is the following mixed integer second order cone program (MISOCP):

$$\min_{(x,y,f^p_{l'})} \quad \sum_{g\in\mathcal{G}}\sum_{t\in\mathcal{T}} C_{g,t}(p_{g,t}) \tag{4.9a}$$

$$s.t. \quad (4.3b)-(4.3g),(4.3j)-(4.3k) \tag{4.9b}$$

$$F^{SOC}_{i,t}(x^D_t)=0, \qquad \forall i\in\mathcal{DB},\forall t\in\mathcal{T} \tag{4.9c}$$

$$(c_t,s_t,f^p_{\mathcal{DL},t},f^q_t)\in\mathcal{OC}^{SOC}, \qquad \forall t\in\mathcal{T} \tag{4.9d}$$

From this problem, we deduce the activation variables y, denoted $\hat{y}$. The second and last step of this approach is to fix the binary variables $\hat{y}$ and solve the following SOCP:

$$\min_{(x,f^p_{l'})} \quad \sum_{g\in\mathcal{G}}\sum_{t\in\mathcal{T}} C_{g,t}(p_{g,t}) \tag{4.10a}$$

$$s.t. \quad (4.3b)-(4.3c),(4.3e)-(4.3g),(4.9c)-(4.9d) \tag{4.10b}$$

$$(p_g,\hat{y}_g)\in\mathcal{GC}^{DC}_g, \qquad \forall g\in\mathcal{TG} \tag{4.10c}$$

$$(p_g,q_g,\hat{y}_g)\in\mathcal{GC}_g, \qquad \forall g\in\mathcal{DG} \tag{4.10d}$$

We compute dispatch decisions x as well as LMPs λ from this problem. The steps of the algorithm for this approach are summarized in Algorithm 3 and Figure 4.6. The idea of calculating prices in this way is suggested in [OSH+05].

Algorithm 3 Algorithm of the Relaxation approach

1: **Step 1:** Solve (4.9) to obtain $\hat{y}$.
2: **Step 2:** Fix the y variables to $\hat{y}$ and solve (4.10) to obtain $\hat{x}$ and $\hat{\lambda}$.

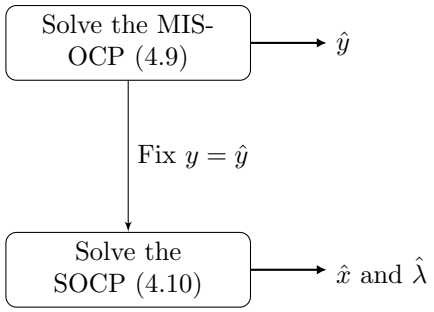


Figure 4.6: Flow chart of the Relaxation approach.

4.6.2 Approach 2: Benchmark

Since the SOC relaxation may produce infeasible dispatch decisions, we consider a benchmark where the binary decisions are still deduced from the relaxation, whereas the dispatch decisions are computed using the AC-OPF equations. After solving (4.9), we solve the following problem:

$$\min_{(x, f^p_{l'})} \quad \sum_{g\in\mathcal{G}}\sum_{t\in\mathcal{T}} C_{g,t}(p_{g,t}) \tag{4.11a}$$

$$s.t. \quad (4.3b) - (4.3c), (4.3e) - (4.3i), (4.10c) - (4.10d) \tag{4.11b}$$

Problem (4.11) is a continuous non-linear (non-convex) problem (NLP). In practice, this class of problems can be solved using interior point (IP) methods or penalty methods (see the first part of chapter 2). Nevertheless, scale might be an obstacle too high to overcome because solving a large-scale NLP is still a challenge in itself.

The steps of the approach are summarized in Algorithm 4 and Figure 4.7.

Algorithm 4 Algorithm of the Benchmark approach (Approach 2).

1: **Step 1:** Solve (4.9) and get $\hat{y}$.
2: **Step 2:** Fix the y variables to $\hat{y}$ and solve (4.11) to get $\hat{x}$ and $\hat{\lambda}$.

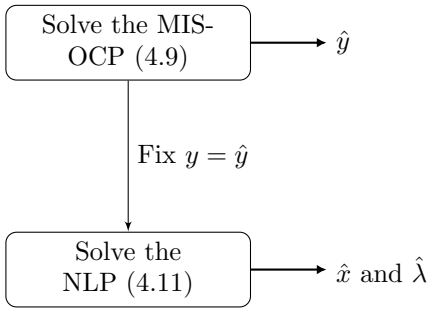


Figure 4.7: Flow chart of the Benchmark approach.

4.6.3 Approach 3: Hybrid Relaxation/Benchmark

To avoid solving large scale NLPs, as is the case in Approach 2, we consider an alternative hybrid approach which combines the ideas of Approach 1 and

Approach 2. In this Hybrid approach, the first two steps are the same as in the Relaxation approach. A step is added in order to recover a feasible dispatch in the distribution network in case the primal solution of the first two steps is infeasible. This approach can drastically reduce the size of the NLPs to solve in the case where several distribution networks are connected to the transmission network. Assuming that we fix the binary decisions $\hat{y}$, the transmission decisions $\hat{x}^T$, as well as the interconnection flow $\hat{f}^p_{l'}$, the distribution problem to solve after the first two steps are computed is the following:

$$\min_{(x^D)} \quad \sum_{g\in\mathcal{DG}}\sum_{t\in\mathcal{T}} C_{g,t}(p_{g,t}) \tag{4.12a}$$

$$s.t. \quad \hat{f}^p_{l',t} = f^{p,D}_{l',t}, \qquad \forall t \in \mathcal{T} \tag{4.12b}$$

$$(4.3\text{h}) - (4.3\text{i}), (4.10\text{d}) \tag{4.12c}$$

Since the SOC relaxation often provides solutions that may be close to AC feasible solutions [BVC20], we rely on the LMPs $\hat{\lambda}$ that are obtained when solving (4.10), even if the dispatch solution x^D may change in the last step of the procedure. This is also motivated by the fact that we do not observe significant differences in prices if one were to compute LMPs with the AC equations. If the prices of (4.10) lead to large LOCs, recovering prices with an AC model should definitely be considered.

The steps of the approach are summarized in Algorithm 5 and Figure 4.8.

Algorithm 5 Algorithm of the Hybrid approach (Approach 3).

1: **Step 1:** Solve (4.9) and get $\hat{y}$.
2: **Step 2:** Fix the y variables to $\hat{y}$ and solve (4.10) to obtain $\hat{x}$ and $\hat{\lambda}$.
3: **Step 3:** If $\hat{x}^D$ is not feasible, solve (4.12) and replace $\hat{x}^D$.

4.6.4 Approach 4: Residual Supply Function

In order to cope with the scale of the problem, we propose a decentralized approach based on distribution network Residual Supply Functions (RSF). This is essentially a hierarchical approach that implements a proxy of a Benders decomposition algorithm in the special case of a convex market clearing problem. The overall spirit of the approach as well as how it should be implemented with the different actors of the market is explained in the example of section 4.3.3. The approach comprises of more steps than the previous approaches, in order to preserve decentralization. We detail the different steps in the following paragraphs.

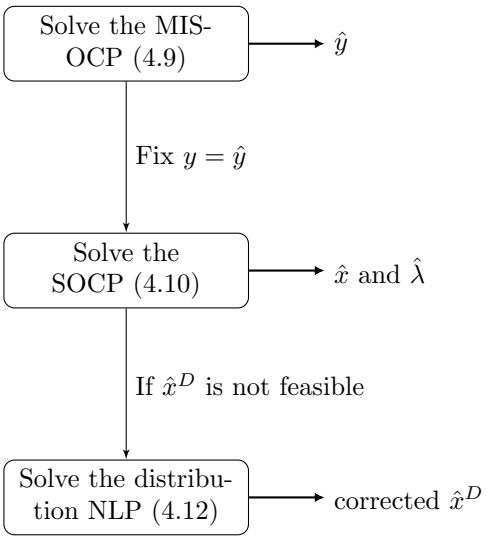


Figure 4.8: Flow chart of the Hybrid approach.

Computing the RSF. As explained in section 4.3.3, the idea of the RSF is to evaluate the cost of exporting real power from the distribution network at the level of the interface. To do so, we solve several distribution problems for different levels of real power flow exports.

Note that, in principle, the RSF is $\mathcal{T}$-dimensional. Indeed, the total cost of export depends on the flow level *for every* time step of the considered horizon. By contrast, real-time energy markets admit one-dimensional bids (a marginal cost curve for each market time unit). In order to overcome this issue, we rely on the fact that the interface flow may not vary drastically from one time step to another. We therefore consider the projection of the total cost function at the vector of equal exports for all time intervals, when computing the RSF.

Concretely, assume an RSF with $N \geq 2$ points. We compute the N points of the RSF for the following levels:

$$E^n_{i'j',t} = E^n_{i'j'} = -S_{i'j'} + \frac{2(n-1)S_{i'j'}}{N-1}, \qquad \forall n \in \{1,\dots,N\}, \forall t \in \mathcal{T}$$

where $S_{i'j'}$ is the line limit of the interface. We see that the quantity defined $E^n_{i'j'}$ is not indexed by time period.

Having decided at which points we discretize the RSF, we now describe how we derive the associated value of the RSF, $\tilde{\mu}_{i',j',t}$, which is time dependent.

Towards this end, we solve the following SOC problem:

$$\mathcal{P}(E^n_{i'j'}): \min \quad \sum_{g\in\mathcal{DG}}\sum_{t\in\mathcal{T}} C_{g,t}(p_{g,t}) \tag{4.13a}$$

$$s.t. \quad (4.3f)-(4.3g),(4.3j),(4.9c)-(4.9d) \tag{4.13b}$$

$$f^p_{i'j',t} = E^n_{i'j'}, \quad (\mu_{i'j',t}) \qquad \forall t\in\mathcal{T} \tag{4.13c}$$

$$y_g \in [0,1]^{d_g}, \qquad \forall g\in\mathcal{DG} \tag{4.13d}$$

This problem is solved using an interior point method, which provides a primal-dual solution. Using a sensitivity argument, the RSF slope $\tilde{\mu}^n_{i'j',t}$ is obtained as the dual optimal value associated with constraint (4.13c) for real power flow level $E^n_{i'j'}$. Note that problem (4.13) relies on a relaxation of the OPF constraints (use of the SOC relaxation) as well as a continuous relaxation of the binary variables y. This is motivated by the fact that having only an approximation of the RSF can provide satisfactory results in practice [PBDS20]. Note that we come up with a more effective way of choosing points on the RSF and this will be illustrated in the results section 4.7.2.

Transmission system dispatch / commitment (primal). Once the RSF is computed, it is explicitly bid into a transmission market clearing model. The way it is bid in our work is through pricing the interconnection line. This pricing is done in the previous step by the ADS and it is assumed that the interconnection line is owned by the DSO. Equivalently, one could think of a virtual BSP placed at the interconnection which bids in the transmission market [PBDS20], [NS20]. To clear the transmission market model, we first solve the following mixed integer linear program (MILP) using the valuation of the interface flow that we compute when deriving the RSF:

$$\min \quad \sum_{g\in\mathcal{TG}}\sum_{t\in\mathcal{T}} C_{g,t}(p_{g,t}) + \sum_{t\in\mathcal{T}}\sum_{n=1}^{N} \tilde{\mu}^n_{i'j',t} f^{p,n}_{i'j',t} \tag{4.14a}$$

$$s.t. \quad (4.3b)-(4.3e),(4.3g) \tag{4.14b}$$

$$f^p_{i'j',t} = \sum_{n=1}^{N} f^{p,n}_{i'j',t}, \qquad \forall t\in\mathcal{T} \tag{4.14c}$$

$$E^n_{i'j',t} \le f^{p,n}_{i'j',t} \le E^n_{i'j',t}, \qquad \forall t\in\mathcal{T}, \forall n\in\{2,\dots,N\} \tag{4.14d}$$

$$y_g \in \{0,1\}^{d_g}, \qquad \forall g\in\mathcal{TG} \tag{4.14e}$$

Note therefore that, in Approach 4, distribution system resources are not bid explicitly into the wholesale market. They are rather aggregated into the RSF, which is computed by the ADS (or DSO), since its derivation requires information about the distribution network.

From this problem, we deduce binary variables $\hat{y}_g, g \in \mathcal{TG}$ for transmission system resources. After fixing the binary variables, we solve the following LP:

$$\min \quad \sum_{g\in\mathcal{TG}}\sum_{t\in\mathcal{T}} C_{g,t}(p_{g,t}) + \sum_{t\in\mathcal{T}}\sum_{n=1}^{N} \tilde{\mu}^n_{i'j',t} f^{p,n}_{i'j',t} \tag{4.15a}$$

$$s.t. \quad (4.3b) - (4.3c), (4.3e), (4.3g), (4.10c), (4.14d) \tag{4.15b}$$

$$f^p_{i'j',t} = \sum_{n=1}^{N} f^{p,n}_{i'j',t}, \quad (\mu_{i'j',t}) \qquad \forall t \in \mathcal{T} \tag{4.15c}$$

From this problem, we deduce dispatch decisions for the transmission network $\hat{x}^T$ (in particular $\hat{f}^{p,T}_{l'} = \hat{f}^p_{l'} = \hat{f}^{p,D}_{l'}$), as well as interface prices $\hat{\mu}$.

Disaggregating interface flows to distribution network commitments and dispatch (primal). Given a target export quantity, the DSO can disaggregate this export level optimally to individual distribution system resources by solving the following MISOCP:

$$\min \quad \sum_{g\in\mathcal{DG}}\sum_{t\in\mathcal{T}} C_{g,t}(p_{g,t}) \tag{4.16a}$$

$$s.t. \quad (4.3j), (4.9c) - (4.9d), (4.12b) \tag{4.16b}$$

$$y_g \in \{0,1\}^{d_g}, \qquad g \in \mathcal{DG} \tag{4.16c}$$

This problem yields distribution binary variables $\hat{y}_g, g \in \mathcal{DG}$ and distribution dispatch decisions $\hat{x}^D$. If the distribution dispatch decisions are not physically implementable, problem (4.12) can be solved in order to ensure feasibility.

Recovering coherent LMPs (dual). Given the solutions of the previous problems, we arrive to a complete binary solution $\hat{y}$ as well as a primal solution $\hat{x}$. The pricing step aims at deriving LMPs λ that are coherent with the primal market clearing solution. We refer to coherent prices as prices that keep the LOC as low as possible. For this purpose, we use the idea exposed in section 4.5.2. Adapting this idea to our context leads to the following transmission network sub-problem, which is an LP:

$$\min \quad \sum_{g\in\mathcal{TG}}\sum_{t\in\mathcal{T}} C_{g,t}(p_{g,t}) + \sum_{t\in\mathcal{T}} \hat{\mu}^n_{i'j',t} f^p_{i'j',t} \tag{4.17a}$$

$$s.t. \quad (4.3b) - (4.3c), (4.3g), (4.10c) \tag{4.17b}$$

Correspondingly, in the distribution network we solve the following SOCP (assuming that the SOCP prices will provide sufficiently accurate prices):

$$\min \quad \sum_{g\in\mathcal{DG}}\sum_{t\in\mathcal{T}} C_{g,t}(p_{g,t}) - \sum_{t\in\mathcal{T}} \hat{\mu}^n_{i'j',t} f^p_{i'j',t} \tag{4.18a}$$

$$s.t. \quad (4.3g), (4.9c) - (4.9d), (4.10d) \tag{4.18b}$$

The LMPs λ are derived as the dual multipliers of constraints (4.3b) in the transmission network, and (4.9c) in the distribution network.

The steps of the algorithm. The steps of the algorithm are summarized in Algorithm 6 and Figure 4.9.

Algorithm 6 Algorithm of the RSF approach (Approach 4).

1: **Step 1:** Compute the RSF by solving (4.13) for N different flow levels.
2: **Step 2:** Solve the transmission system primal problem (4.14) and derive the transmission binary decisions $\hat{y}_g,\ g \in \mathcal{TG}$.
3: **Step 3:** Fix the transmission binary decisions and solve (4.15) in order to derive the transmission dispatch decisions $\hat{x}^T$ and the interface prices $\hat{\mu}$.
4: **Step 4:** Clear the distribution problem (4.16) in order to derive the distribution binary decisions $\hat{y}_g,\ g \in \mathcal{DG}$ and distribution dispatch decisions $\hat{x}^D$.
5: **Step 5:** If $\hat{x}^D$ is not feasible, solve (4.12) and update $\hat{x}^D$.
6: **Step 6:** Solve (4.17) and (4.18) in order to obtain LMPs $\hat{\lambda}$.

4.6.5 Comparison of the Approaches

Parallelization In general, a transmission network is connected to numerous distribution networks. In such a setting, the Hybrid and RSF approach can be parallelized. Indeed, step 1 of Algorithm 5 can be parallelized by solving (4.12) for each distribution network in parallel. Despite the fact that this step can be performed in parallel, the overall performance of Algorithm 5 would not improve significantly because most of the computational effort is put on the first steps of the algorithm. Therefore, we will not consider a parallel version of this approach.

On the contrary, a number of steps of the RSF approach (Algorithm 6) can be executed in parallel. Step 1 of the RSF can be the parallelized at two points: each distribution network can compute the RSF independently, and each of the N flow levels can be computed as an independent problem. Using the natural decomposability of the network, Steps 4, 5, and 6 can also be executed in parallel.

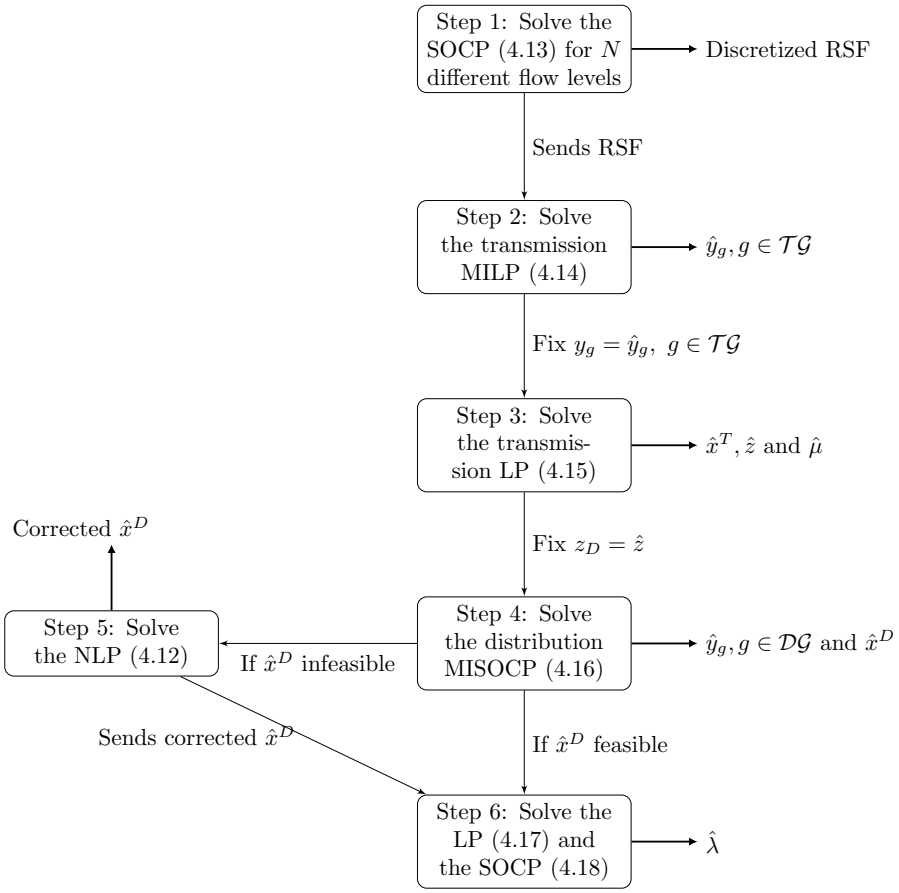


Figure 4.9: Flow chart of the RSF approach.

Relative Merits of Each Approach A list of strengths and weaknesses of each approach is presented in Table 4.3.

4.7 Numerical Illustration

We first detail the data that we used for the numerical simulations (section 4.7.1.1) and we experimentally explain the setting chosen for the RSF approach (section 4.7.1.2 and section 4.7.1.3). In section 4.7.2, we compare the approaches implemented in this chapter on four realistic test systems.

Table 4.3: Comparison of the advantages and drawbacks of the proposed approaches.

Approach	Advantages	Drawbacks
Relaxation	• Fastest way to derive a relatively accurate relaxed solution.	• Centralized market clearing: one operator responsible for the complete network. • Potentially infeasible distribution dispatch.
Benchmark	• Ensures feasible dispatch. • Local optimum guarantee (if solved).	• Centralized market clearing: one operator responsible for the complete network. • Requires solving a large-scale NLP (T&D network).
Hybrid	• Ensures feasible dispatch. • Parallelizable.	• Centralized market clearing: one operator responsible for the complete network.
RSF	• Decentralized market clearing: TSO & DSO only share interconnection information. • Highly parallelizable. • Ensures feasible dispatch.	• Depends on the precision of the approximated RSF. • Heuristic solution approach. • Primal solution can be far from optimal.

4.7.1 Data and Parametrization of the RSF Approach

4.7.1.1 Test Cases

The test cases that we consider are derived from data sets that were used in the European project SmartNet (`http://smartnet-project.eu/`). For each of the test cases that we present, we receive as input the topology of the T&D network, the bids associated with the generators at each node, and a time horizon of 3 or 4 time steps. Given that each market time unit of the European balancing market corresponds to a 15-minute step, the horizon of the problem corresponds to 45 minutes or 1 hour.

We consider networks from the Italian and Danish power systems. For the Italian system, we consider three test cases: a medium-sized one, called `small_it`, and two other ones based on the same network topology, called `Italy1` and `Italy2`. For the Danish system, we consider only one test case called `Denmark`. Instance `small_it` is a medium-sized example, which serves towards validating our approaches, while `Italy1`, `Italy2` and `Denmark` are more realistic instances. An overview of the Italian and Danish test systems is provided in Table 4.4. Table 4.4 reports for each test case: the number of time-steps, the number of transmission buses, the number of distribution buses, the number of distribution networks, the number of bids, the transmission network real power imbalance (TNRPI) in megawatt (MW), the distribution network

RPI (DNRPI), the total RPI (TRPI), the transmission network real power volume bid (TNRPVB), the distribution network RPVB (DNRPVB), the total RPVB (TRPVB), the number of binary variables, the number of variables and the number of constraints. RPI and RPVB are obtained as follows:

$$\text{TNRPI} = \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{TB}} D^p_{i,t},\ \text{DNRPI} = \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{DB}} D^p_{i,t},\ \text{TRPI} = \text{TNRPI} + \text{DNRPI},$$

$$\text{TNRPVB} = \left[\sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{TG}} \min(0, \underline{p}_{g,t}); \sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{TG}} \max(0, \overline{p}_{g,t})\right],$$

$$\text{DNRPVB} = \left[\sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{DG}} \min(0, \underline{p}_{g,t}); \sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{DG}} \max(0, \overline{p}_{g,t})\right],$$

$$\text{TRPVB} = \left[\sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}} \min(0, \underline{p}_{g,t}); \sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}} \max(0, \overline{p}_{g,t})\right].$$

Table 4.4: Overview of the Italian and Danish test cases used in the numerical experiments.

Test Case	`small_it`	`Italy1`	`Italy2`	`Denmark`
$\lvert\mathcal{T}\rvert$	4	3	3	4
$\lvert\mathcal{TB}\rvert$	27	4,236	4,236	209
$\lvert\mathcal{DB}\rvert$	175	1,822	1,822	2,981
# DNs	4	50	50	73
# Bids	1,667	12,318	26,578	25,923
TNRPI (MW)	-88.98	+226.67	-1,170.16	-13.85
DNRPI (MW)	+220.47	-526.77	-209.58	-138.95
TRPI (MW)	+131.48	-300.10	-1,379.74	-152.80
TNRPVB (MW)	$[-3,429\ ;\ 24]$	$[-50,379\ ;\ 69,637]$	$[-61,459\ ;\ 63,388]$	$[-12,028\ ;\ 13,420]$
DNRPVB (MW)	$[-449\ ;\ 36]$	$[-82\ ;\ 47]$	$[-443\ ;\ 139]$	$[-955\ ;\ 353]$
TRPVB (MW)	$[-3,877\ ;\ 60]$	$[-50,460\ ;\ 69,684]$	$[-61,902\ ;\ 63,527]$	$[-12,983\ ;\ 13,774]$
# Binary	4,222	29,660	51,510	90,688
# Variables	25,342	244,052	298,226	391,040
# Constraints	26,401	248,864	310,272	429,931

4.7.1.2 Enhancing the Computation of the RSF

As explained in the first paragraph of section 4.6.4, the RSF is computed by using equally spaced points in the interval imposed by the line limits of each interface line. Since these limits are often large compared to the potential values of the interface flow, taking equally spaced points might cause computing the RSF on flow values that are far from being optimal. As an example,

Figure 4.10 shows that the computation of the RSF on the import side (negative x values) lacks of interest and one could assume that importing, in that case, is costless. Ideally, one should detect where it is relevant to add points to the RSF. Assuming that we decide to build the RSF with N points, what we suggest is the following: (i) evaluate the RSF on $N/2$ equally spaced points between lower and upper interface flow limits; (ii) detect the 3 consecutive points with the highest price difference, and denote their flow values f_0, f_1, f_2; (iii) evaluate the RSF on $N/2$ equally spaced points between f_0 and f_2. By applying this, we see how the precision of the blue curve improves compared to the equally spaced points curve in magenta (Figures 4.10 and 4.11 and in particular 4.11(b) compared to 4.10).

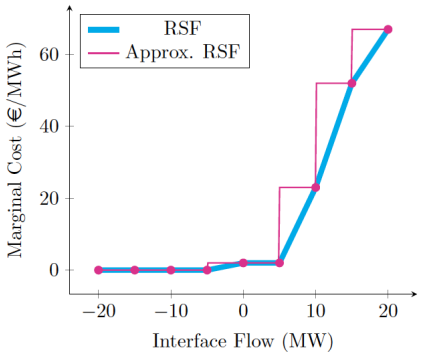
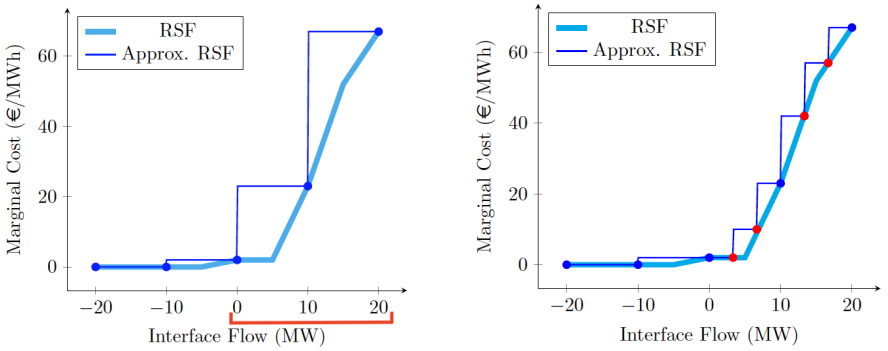


Figure 4.10: Example of the basic computation of the RSF.

To measure the impact of the precision of the RSF, we report the evolution of the primal objective values as well as the LOCs for the medium-size example, `small_it`, for 10, 20 and 30 points on the RSF (Figure 4.12). The figures show how the enhancement step increases the quality of the solution with the same number of points on the RSF. Even if the enhancement makes the approach less parallelizable (one needs to split the computation for the first $N/2$ points before computing the remaining $N/2$ as opposed to directly compute N points), we will use this enhancement when displaying the results in section 4.7.2.

4.7.1.3 Number of Points on the RSF

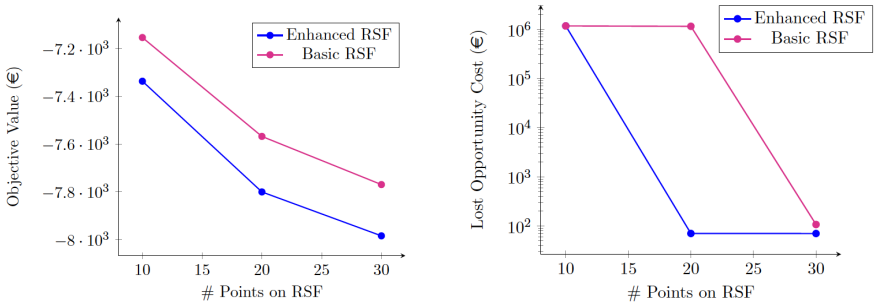
We conduct a sensitivity analysis on the number of points that should be considered when computing the RSF (Step 1 of Algortihm 6). To do so, we report the objective value, the lost opportunity cost and the solve time for a different number of points on the RSF for the test case `Italy1` (Figure 4.13). The general trend is as expected: the solution improves as the number of points increases. Note however that the objective value and the LOC are not strictly



(a) Step (i): Evaluate the RSF for -10 MW, 0 MW, 10MW, 20 MW. Step (ii): detect the points of interest: 0, 10 and 20 MW.

(b) Step (iii): Evaluate the RSF for 4 additional points between 0 and 20 MW.

Figure 4.11: Enhancing the RSF by intelligently choosing the points to compute.

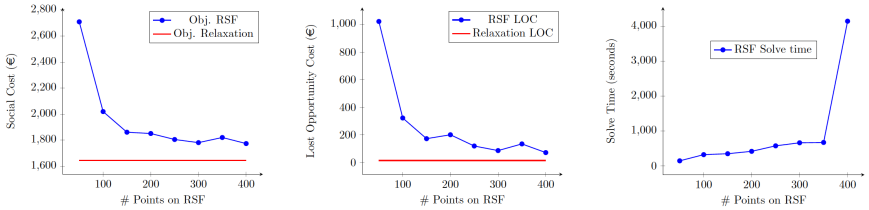


(a) Comparison of the objective value.

(b) Comparison of the LOC.

Figure 4.12: Evolution of the objective value and the LOC of the `small_it` test case as a function of the number of points on the RSF by using two different point selection strategies: Basic vs. Enhanced.

decreasing as a function of the number of points: indeed, this is due to the discretization of the RSF that can be favorable even if fewer points are used. It should also be noted that considering 400 points for this particular test case leads to a costly computation (+4,000 seconds for this run). Even if the approach is parallelizable, the additional computational cost of having 400 points appears to be vain compared to the solutions with 300 or 350 points. For this reason, in the simulations, we decide to adopt 300 points on the RSF for `Italy1`, `Italy2` and `Denmark` which are of the same scale. Since `small_it` is



(a) Objective value. (b) LOC. (c) Solve time.

Figure 4.13: Assessment of the RSF approach as a function of the number of points considered in Step 1 of Algortihm 6.

of smaller scale, we consider only 100 points for this test case.

4.7.2 Simulations

Computational setting. The clearing algorithms are implemented in Julia (version 1.1.1) using JuMP on a MacBook Pro 2016, with a 2.9 GHz Dual-Core Intel Core i5 processors. LPs, SOCPs, MILPs, MISOCPs, are solved using Mosek (version 9); NLPs are solved using IPOPT (version 3.13.2).

Tolerance. We set the feasibility tolerance of the solvers to $\epsilon = 1e^{-6}$. When showing the results, the only constraint that might be violated is (1.9a) because it is relaxed as (1.13) in Approach 1 (Relaxation). The maximum constraint violation is then reported in Table 4.5 in column Max. Viol..

Numerical results. Each approach is tested on each test case. The results are presented in Table 4.5. From left to right, the columns display the name of the test case, the approach used, the objective value of the primal problem, the absolute gap to the lower bound (obtained using the Relaxation approach), the maximum violation of a constraint of the primal solution, the PLP, the LOC, and the solve time in seconds. When reporting the RSF approach, we accompany it with the number of points chosen to compute the RSF (100 for `small_it`, 300 for the other test systems).

The results on the medium-sized example, `small_it`, demonstrate that all the approaches provide similar results in terms of objective value. For the largest instances (`Italy1`, `Italy2` and `Denmark`), the objective values of the Relaxation and Hybrid approaches confirm the quality of the SOCP relaxation of the AC-OPF on radial networks. Note, however, that for the Danish test system, even if the Relaxation solution is almost feasible ($MV = 2^{-6}$), both Hybrid and RSF provide solutions for which the objective is significantly different. We further note that the solve time for all the approaches is significant for

Table 4.5: Results of the different approaches on the set of SmartNet test cases.

Test case	Approach	Objective	Gap (€)	Max. Viol.	PLP (€)	LOC (€)	Time (s)
`small_it`	Relaxation	$-7.872e^{+3}$	-	$1e^{-6}$	$3.559e^{+5}$	0.59	5.39
	Benchmark	$-7.870e^{+3}$	2.25	$1e^{-9}$	$3.559e^{+5}$	0.16	17.03
	Hybrid	$-7.872e^{+3}$	0.01	$3e^{-7}$	$3.559e^{+5}$	0.53	10.17
	RSF-100	$-7.864e^{+3}$	8.09	$2e^{-7}$	$3.561e^{+5}$	109.88	85.90
`Italy1`	Relaxation	$1.644e^{+3}$	-	$2e^{-4}$	$2.195e^{+6}$	14.66	85.54
	Benchmark	-	-	-	-	-	-
	Hybrid	$1.648e^{+3}$	3.79	$5e^{-7}$	$2.195e^{+6}$	14.30	89.23
	RSF-300	$1.781e^{+3}$	136.52	$9e^{-7}$	$2.219e^{+6}$	86.37	662.5
`Italy2`	Relaxation	$7.998e^{+3}$	-	$1e^{-4}$	$2.825e^{+6}$	29.68	160.7
	Benchmark	-	-	-	-	-	-
	Hybrid	$8.001e^{+3}$	2.57	$4e^{-7}$	$2.825e^{+6}$	32.25	166.9
	RSF-300	$8.147e^{+3}$	149.22	$1e^{-6}$	$2.852e^{+6}$	166.01	823.2
`Denmark`	Relaxation	$-1.023e^{+4}$	-	$2e^{-6}$	$1.633e^{+5}$	30.30	362.6
	Benchmark	-	-	-	-	-	-
	Hybrid	$-9.951e^{+3}$	274.92	$8e^{-7}$	$1.633e^{+5}$	33.41	375.4
	RSF-300	$-9.518e^{+3}$	707.97	$9e^{-7}$	$1.670e^{+5}$	592.57	1,340

the Danish test case (more than 5 minutes). Executing the RSF sequentially also leads to time-consuming computations. Fig. 4.14 presents how parallelization can decrease substantially the execution time of the RSF approach. In particular, using 16 processors ensures execution times of less than 5 minutes for the three largest test cases. On the contrary, it is not possible to parallelize centralized schemes. This underlines the potential weaknesses of considering centralized schemes in addition to not preserving privacy.

Insofar as the RSF approach is concerned, the optimality guarantee worsens and could be improved with the number of points chosen. When analyzing the solution of the RSF approach compared to the Benchmark approach, we notice that in the Italian test cases (`Italy1` and `Italy2`) the RSF approach underestimates what can be withdrawn from the distribution network. In addition to underestimating what the distribution network could cover in terms of power balance, one of the assumptions on which we rely is partially wrong: indeed, assuming that real power flows at the interface are not changing from one time-step to another can lead to bad approximations of the RSF. Consequently, some decisions look drastically different. Nevertheless, the adaptability of the RSF can lead to consider building the RSF around a predetermined dispatch instead of assuming the discretization between the two bounds of the line limits. This is not explored in this work.

That being said, the differences between objective values of methods ensuring a feasible dispatch (Hybrid and RSF) are quite small. Moreover, compared to the trading volume of the market (an estimation is for example PLP), these differences could even be assumed negligible as well as the LOC which is always larger when using the RSF approach than the centralized approaches.

To sum up, the RSF approach provides slightly suboptimal solutions while

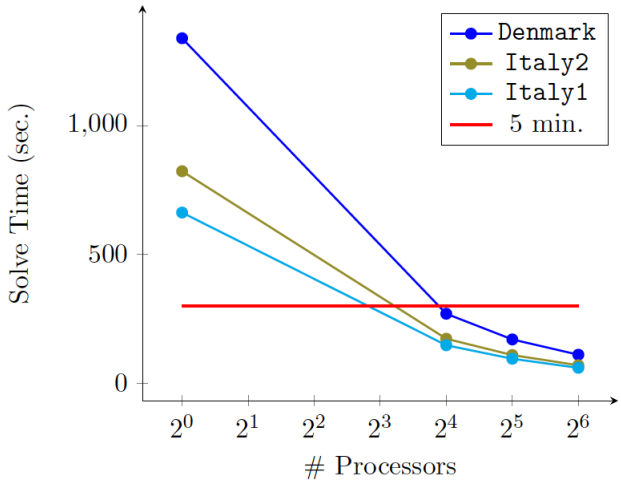


Figure 4.14: Evolution of the solve time of the RSF approach with respect to the number of processors on the Italian and Danish test cases.

maintaining privacy of the different actors and relies on decomposability which allows for heavy parallelization of the framework. This validation of the RSF approach is not only shown on a medium-sized example, `small_it`, but also on three national-scale realistic instances of the problem (`Italy1`, `Italy2`, `Denmark`).

4.7.3 Settlements in the Danish Case Study

Moreover, the RSF approach we develop implements a simple market design aiming at fairly remunerating each actor participating in the market. Practically, we draw a settlement table to show how the approach is able to define the cash flows and remunerations of each participant. We illustrate the insights of such settlement tables on the Danish test system (Table 4.6).

Table 4.6: Settlement table of the `Denmark` test case.

	Transmission			Distribution		
Settlements	BSPT	BRPT	TSO	ADS	BSPD	BRPD
TM-BSP	+11,371 €	-	−8,311 €	−3,061 €	-	-
ADS Dis.	-	-	-	−568 €	+568 €	-
TM-BRP	-	−803,062 €	+808,713 €	-	-	−5,651 €
ADS Rebal.	-	-	−5,651 €	+5,651 €	-	-
Total	−11,371 €	−803,062 €	+794,751 €	+2,022 €	+568 €	−5,651 €

In Table 4.6, we aggregate the transmission (resp. distribution) BSPs

(BSPT resp. BRPD), BRPs (BRPT resp. BRPD) and the total remunerations or payments are reported (over the 4 time-steps of the test case) in the same way as the example of section 4.3.3. In this test case, the power is mostly flowing from the transmission network to the different distribution networks: the ADS is buying power from the transmission network for a total cost of 3,061€. The ADS compensates the distribution BSPs when disaggregating. The TSO is collecting payments from BRPs in both transmission and distribution before reallocating the BRPD payments to the ADS. In total, the TSO and the ADS are collecting congestion revenue, the TSO's revenue being significantly more important.

Even if the number of buses is more important in the distribution network in this test case, we observe that most of the power is consumed and produced in the transmission network (Table 4.7), which explains the scale of the revenues and costs.

Table 4.7: Power produced and consumed in the `Denmark` test case.

Transmission		Distribution	
Power injected	Power withdrawn	Power injected	Power withdrawn
19,756 MW	-19,561 MW	550 MW	-723 MW

To understand why the TSO is collecting an important congestion revenue, we draw a box plot representation of the transmission and distribution LMPs computed for one time-step ($t = 89$[1]) on the `Denmark` test case in Figure 4.15. When comparing the LMPs of both networks, we notice that even if the lower, middle and upper quartiles are of the same scale, the spread of the transmission LMPs is more important. This observation stresses the potential significant differences between prices at certain locations. Note also some negative transmission prices: 208 MW withdrawn for $t = 89$ have negative valuations, i.e. some consumers are paid to consume power instead of paying which leads to negative prices in certain locations where only this type of bid has been accepted. The disparity in locational pricing implicitly shows how congested the transmission network is.

4.8 Conclusions & Perspectives

In this chapter, we develop and implement a decentralized market clearing T&D platform capable of respecting several technical requirements. The platform offers several advantages:

(i) it preserves the privacy of the TSO and the DSOs by a careful exchange of only border information at the interface.

[1] In this test case, $\mathcal{T} = \{89, 90, 91, 92\}$.

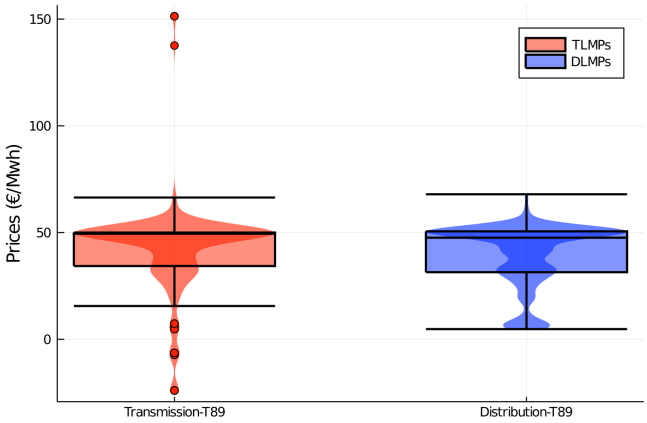


Figure 4.15: A box plot representation of the LMPs in the Danish test case for $t = 89$.

(ii) dispatch decisions are AC feasible in the distribution network, which is relevant due to distribution system flexibility and renewable supply.

(iii) in addition to primal decisions, the LMPs cleared in a straightforward way provide satisfactory results and incentivize the use of flexibility and signal investments in specific locations.

(iv) the RSF approach is highly parallelizable and respects the time-limit imposed by short-term markets.

(v) by comparing the RSF approach to three other centralized approaches, impossible to apply in practice, we show the slight suboptimality of the RSF approach.

(vi) intensive testing on large-scale systems shows the applicability of the RSF approach in a realistic setting.

These aspects stress the benefit of the prototyped platform implemented in this chapter.

Even if the platform is already promising, improvements are already worth considering. For example, the selection of points of interest on the RSF could largely be improved through market experience and historical data. To validate this work, broader experiments on different topologies should be tested. Also, the LMPs were cleared using IP pricing [OSH$^+$05] and a comparison of different price clearing techniques are of interest to enhance the quality of the framework. Extending [GNB20] by considering large-scale test systems to assess the pricing techniques could definitely be studied in the future. The RSF approach could also be applied towards hierarchical balancing in markets with zonal pricing

[PBDS20]. The goal is to explore the analysis of that chapter to a realistic model of the Nordic system.

Appendix

4.A Details on the Bid Structure

The market clearing model presented in this chapter is inspired by the products that are available in the Central Western European (CWE) day-ahead energy market. To do so, bids are modeled through market orders as suggested by EUPHEMIA[2]. The first unit that we will use in order to define a complete bid is the segment bid. A segment bid (or S-bid) is characterized by a minimum and maximum quantity of real power and a certain marginal cost. Note that we allow consumption and production bids, so we have no assumptions on the signs of the minimum and maximum quantity. A Q-bid links several segment bids. We make explicit the relationships between segment bids when defining bid constraints. We also link Q-bids over time and we refer to such bids as Qt-bids. A bid is associated with a certain BSP g, at a certain moment t.

Then, a segment bid (g, t, qt, q, s) is defined by the following 5 fields [MRS+17]:

1. a BSP $g \in \mathcal{G}$, located at a certain bus $i \in \mathcal{B}$,
2. a time-step $t \in \mathcal{T}$,
3. a Qt field (or Qt-bid) $qt \in QtB$,
4. a Q field (or Q-bid) $q \in QB$, and
5. a segment (or S-bid) $s \in SB$.

This segment bid is associated to the Q-bid (g, t, qt, q) which is associated to the Qt-bid (g, t, qt). To make the explanations more concrete, the reader can refer to the example on Figure 4.16. Each bid can be rejected, partially accepted or totally accepted by an operator maximizing the welfare (or minimizing the total activation cost) over the transmission and distribution network. Each S-bid sb is associated with a cost $c_{sb}(x_{sb}) = a_{sb}(P_{sb}x_{sb})^2 + b_{sb}P_{sb}x_{sb} + c_{sb}$ (it is the general form but in the dissertation, this cost is linear). Here, x_{sb} is the fraction of acceptance of the bid and P_{sb} is the difference between the maximum and minimum quantity of the bid.

Bids are associated with specific attributes that provide a rich set of options for distributed resources to represent complex operating constraints for their assets. The constraints we consider in the dissertation to expand $\mathcal{GC}$ are the

[2] https://www.nordpoolgroup.com/globalassets/download-center/single-day-ahead-coupling/euphemia-public-description.pdf

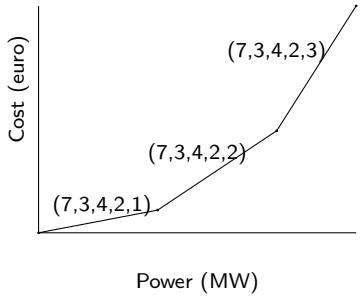


Figure 4.16: Example of a bid. $g = 7$ stands for the BSP, $t = 3$ for the time-step, $qt = 4$ for the Qt-field, $q = 2$ for the Q-field. There are 3 S-bids (7,3,4,2,1), (7,3,4,2,2), (7,3,4,2,3) associated to the Q-bid (7,3,4,2). The Q-bid (7,3,4,2) is part of the Qt-bid (7,3,4).

following:

$$p_{g,t} = \sum_{sb=(g,t,qt,q,s)} P_{sb} x_{sb},\ \forall g \in \mathcal{G},\ t \in \mathcal{T} \tag{4.19}$$

$$\underline{x}_{sb} s_{sb} \leq x_{sb} \leq s_{sb} \overline{x}_{sb},\ \forall sb \in SB \tag{4.20}$$

$$s_{(g,t,qt,q,s)} \leq q_{(g,t,qt,q)},\ \forall (g,t,qt,q,s) \in SB \tag{4.21}$$

$$q_{(g,t,qt,q)} \leq qt_{(g,t,qt)},\ \forall (g,t,qt,q) \in QB, (g,t,qt) \in QtB \tag{4.22}$$

$$qt_{(g,t,qt)} \leq \sum_{qb=(g,t,qt,q)} q_{qb},\ \forall (g,t,qt) \in QtB \tag{4.23}$$

$$q_{(g,t,qt,q)} - q_{(g,t-1,qt,q)} - \alpha_{(g,t,qt,q)} + \omega_{(g,tq,t,q)} = 0,\ \forall (g,t,qt,q), (g,t-1,qt,q) \in QB \tag{4.24}$$

$$\alpha_{qb} + \omega_{qb} \leq 1,\ \forall qb \in QB \tag{4.25}$$

$$q_{(g,t,qt,q)} \geq \alpha_{(g,\tau,qt,q)}, \forall ((g,t,qt,q),(g,\tau,qt,q)) \in MDP \tag{4.26}$$

$$\sum_{qb \in exqb} q_{qb} \leq 1,\ \forall exqb \in ExQB \tag{4.27}$$

$$\sum_{qtb \in exqtb} qt_{qtb} \leq 1,\ \forall exqtb \in ExQtB \tag{4.28}$$

$$\underline{RP}_{g,t} \leq p_{g,t+1} - p_{g,t} \leq \overline{RP}_{g,t},\ \forall g \in \mathcal{G}, t \in \mathcal{T} - \{t_f\} \tag{4.29}$$

$$0 \leq x_{sb} \leq 1,\ sb \in SB \tag{4.30}$$

$$s \in \{0,1\}^{|SB|}, q \in \{0,1\}^{|QB|}, qt \in \{0,1\}^{|QtB|}, \alpha, \omega \in \{0,1\}^{|QB|} \tag{4.31}$$

(4.19) describes how bids impact the net injection of real power. (4.20) defines the activation of a segment bid. (4.21) imposes that a segment-bid is activated

only if the associated Q-bid is also activated. The same holds with a Q-bid and the associated Qt-bid in (4.22). (4.23) ensures that a Qt-bid is activated if at least one of the associated Q-bids is also activated. (4.24) defines that two consecutive Q-bids of a Qt-bid are linked: the Q-bid at t can only be activated if the one at $t-1$ has also been activated. Constraint (4.25) imposes the fact that a bid cannot be starting and ending at the same time. (4.26) is ensuring that if a bid is activated, it remains active for a minimum amount of time. (4.27), (4.28) indicate that certain Q-bids or Qt-bids should be activated only if others are not (i.e. an exclusive choice has to be made). (4.29) are ramp constraints on real power outputs. (4.30) and (4.31) denote that x should be fractional as opposed to the other bid-related variables which are binary variables.

The collection of these constraints are implemented in the T&D market-clearing platform of the chapter.

Chapter 5

A Game-Theoretical Analysis of T&D Coordination Schemes

5.1 Introduction

The current paradigm of power system operations mostly focuses on the transmission network. Given the vast amount of unexploited flexible resources that are connected to the distribution network, the existing power system paradigm puts an important part of the system aside.

The transmission network is the only part of the electricity supply chain that is currently optimized. The flexibility in the distribution network is mainly originating from active residential and commercial demand-side management, which we will need to exploit effectively in the coming decades if we wish to maintain the quality of service that we currently enjoy [CH11]. However, the distribution network is, in itself, a system of massive scale which presents a host of operational challenges. The amount of renewable resources that are located in the distribution network, mainly in the form of solar panels, has been growing and becoming an increasingly important component of the electric power supply chain. Due to distribution constraints and the unpredictability of renewable resources, a certain amount of this renewable power needs to be consumed locally [fSG14]. Coordination of operations in electricity markets has also been discussed in [SOAS10, KZ13, YH17].

In this chapter, we use the concept of Generalized Nash Equilibrium (GNE) to model certain TSO-DSO coordination schemes introduced in chapter 3. Using GNEs to model interactions and designs in electricity markets is common practice, specifically for TSO coordination [SOAS10], market design [EN09,

LCJWA20] or competition between generators [JYS99, Hob01]. Even if they are appealing from a modeling point of view, Generalized Nash Equilibrium Problems (GNEPs) are difficult problems to tackle (see section 5.3 for more details).

Part of this work draws inspiration from [SOAS10], where the authors focus on the counter-trading of re-dispatching resources between two Transmission System Operators (TSOs), in the context of congestion management. The authors investigate whether there should exist a separate market for transmission capacity by resorting to GNE, due to the influence of each TSO's action on the other TSO's decisions. The authors rely on GNE and not simply Nash Equilibrium (NE) because of the coupling constraints arising at interconnections between TSOs. This modeling is particularly interesting because it gives the opportunity to highlight potential asymmetric valuations of border flows. In this setting, it is possible to span multiple equilibria and discuss their efficiency as well as assess the quality the coordination schemes considered.

We transpose the framework of [SOAS10] to the context of TSO-DSO coordination, where the activation of distribution system reserves by the TSO has an impact on the feasible actions of DSOs. We specifically focus on two coordination schemes inspired by the EU SmartNet project on TSO-DSO coordination [GPS18,MRS+17]. The coordination schemes have been introduced in Chapter 3 and the two coordination schemes considered are Shared Balancing Responsibilities (SBR) and Local Ancillary Services (LAS) Markets. In the remainder of the chapter, we interpret SBR as a Generalized Nash Equilibrium Problem (GNEP). We envision two different alternatives for LAS. Indeed, we develop one alternative which relies on a non-cooperative simultaneous game, where the TSO has limited access (TLA) to DSO resources. The other interpretation for this scheme relies on a hierarchical configuration, in which we assume that the DSO solves its local imbalance first before the TSO clears the rest of the market. This assumption (on DSO solving local imbalance) contradicts the interpretation of chapter 4 and most of common practices in current European markets but we still envision such a scheme since the role of the DSO is evolving. We name this scheme Local Markets (LM). Such a sequential interaction involves a hierarchical interaction that can be represented as a Stackelberg game (leader-follower model) under the assumption that the leader, the DSO, anticipates the rational reaction of the follower, the TSO [DD12,LC19]. In this case, the leader incorporates explicitly in its optimization problem the rational reaction function of the follower. The closed form expression of the latter is obtained by solving first the follower's optimization problem at the lower level of the Stackelberg game, considering as fixed the decision variables of the leader. The leader, at the upper level, then incorporates the follower's rational reaction function, expressed as a function of the leader's decision variables only, directly in its optimization problem, thereby proceeding backwards.

The focus of our chapter is (i) to model various TSO-DSO coordination

schemes which have been proposed in the SmartNet project as non-cooperative games or Stackelberg games, (ii) to interpret the solutions, and (iii) to compare the relative strengths and weaknesses of the different schemes and in particular possible allocation inefficiencies.

The remainder of the chapter is organized as follows: we briefly remind the context of TSO-DSO coordination in section 5.2. We provide an introduction to game theoretical concepts used in this chapter in section 5.3. We present the models of the TSO-DSO coordination schemes in section 5.4. The different schemes are illustrated through numerical results presented on a toy example in section 5.5. Section 5.6 concludes the chapter.

5.2 General Assumptions and Market Structure

5.2.1 Topology of the Network

In this chapter, we follow the same assumptions as in the previous chapters. We consider a meshed transmission network, on which we apply the DC approximation of the power flow equations, connected to several distribution networks. The distribution networks are radial and we apply the SOCP relaxation on these networks. We do not consider recovering AC feasible dispatches in this chapter. The model presented will be considering only one distribution network connected to the transmission network for simplicity (Figure 3.1). We consider a power network $\mathcal{N} = (\mathcal{B}, \mathcal{L})$ where $\mathcal{B} = \mathcal{TB} \cup \mathcal{DB}$ is the set of buses (transmission and distribution) and $\mathcal{L} = \mathcal{TL} \cup \{(i', j')\} \cup \mathcal{DL}$ is the set of lines (transmission, interconnection and distribution). The centralized optimization

problem to solve is then:

$$\min \quad \sum_{g\in\mathcal{G}} C_g(p_g) \tag{5.1a}$$

$$s.t. \quad \sum_{g\in\mathcal{G}(i)} p_g - D_i^p = \sum_{j=\delta(i)} f_{ij}^{p,T}, \qquad \forall i \in \mathcal{TB} \tag{5.1b}$$

$$f_{ij}^{p,T} = B_{ij}(\theta_i - \theta_j), \qquad \forall (i,j) \in \mathcal{TL} \cup \mathcal{TL}^R \tag{5.1c}$$

$$(p_{\mathcal{TG}}, f_{\mathcal{TL}}^{p,T}, \theta) \in \mathcal{C}^T \tag{5.1d}$$

$$\sum_{g\in\mathcal{G}(i)} p_g - D_i^p = \sum_{j=\delta(i)} f_{ij}^{p,D} + G_i c_{ii}, \qquad \forall i \in \mathcal{DB} \tag{5.1e}$$

$$\sum_{g\in\mathcal{G}(i)} q_g - D_i^q = \sum_{j=\delta(i)} f_{ij}^{q} - B_i c_{ii}, \qquad \forall i \in \mathcal{DB} \tag{5.1f}$$

$$f_{ij}^{p,D} = -G_{ij}c_{ii} + G_{ij}c_{ij} - B_{ij}s_{ij}, \qquad \forall (i,j) \in \mathcal{DL} \cup \mathcal{DL}^R \tag{5.1g}$$

$$f_{ij}^{q} = B_{ij}c_{ii} - G_{ij}s_{ij} - B_{ij}c_{ij}, \qquad \forall (i,j) \in \mathcal{DL} \cup \mathcal{DL}^R \tag{5.1h}$$

$$(p_{\mathcal{DG}}, q, f_{\mathcal{DL}}^{p,T}, f^q, c, s) \in \mathcal{C}^D \tag{5.1i}$$

$$f_{i'j'}^{p,T} = f_{i'j'}^{p,D} \tag{5.1j}$$

$$-S_{i'j'} \leq f_{i'j'}^{p,T} \leq S_{i'j'} \tag{5.1k}$$

$$-S_{i'j'} \leq f_{i'j'}^{p,D} \leq S_{i'j'} \tag{5.1l}$$

The transmission constraints are the real power balance constraints (5.1b), the real power flow definition (5.1c) and an aggregated version of the other classical inequality constraints (5.1d) (generation capacity, line capacity etc.). The distribution constraints follow the same spirit: (5.1e)–(5.1f) represent the power balance constraints, (5.1g)–(5.1h) show the flow definitions and the engineering constraints are cast in (5.1i). The interconnection real power flow is duplicated using $f_{i'j'}^{p,T}$ for the transmission network and $f_{i'j'}^{p,D}$ for the distribution network. Constraint (5.1j) ensures the equality of the duplicated variable. Interconnection flow has a specific capacity in constraints (5.1k) (5.1l).

5.2.2 Market Structure

In this work, we focus on the interaction of a TSO with a group of DSOs in the real-time market, where reserves are activated in order to balance potential load disturbances. The decisions of each operator can be seen as adjustments on a predetermined dispatch.

We also assume that adjusting power generation in the distribution network is cheaper than in the transmission network. Indeed, the transmission network includes power plants with non-negligible fuel costs, and adjusting

power generation in these resources might be more costly. We assume that adjusting generation in the distribution network can be achieved through demand response or local renewable resources, which we considered to be cheap adjustment actions for the sake of developing the discussion in the present chapter. However, losses cannot be neglected in the distribution network, and even if the marginal activation cost of a distribution system reserve is cheaper, the flow of power in the distribution network might increase the cost of activation. Currently, the use of distributed resources is limited and uncoordinated operations can then lead to highly inefficient outcomes in terms of welfare. These assumptions and the schemes that we propose in this chapter are in line with what can be found at the EU level in [GPS18].

In a decentralized market structure, each DSO resolves local grid issues at the lowest possible cost. The role of the TSO, which ignores distribution system constraints in order to allow for scalability of operations, can vary according to the type of coordination scheme that we consider. In this chapter, we consider three different configurations:

(i) The TSO and the DSO only rely on the resources connected on their own network (shared balancing responsibilities).

(ii) The TSO has partial access to distribution system resources (TSO has limited access to DSO resources).

(iii) The DSO anticipates the local imbalances, and makes the remaining distributed generation capacity available to the TSO (local markets).

These schemes are compared to the centralized optimal solution, which is the benchmark and provides the most efficient solution.

5.3 Preliminaries on Game Theory

In this chapter, we model two schemes as non-cooperative simultaneous games (SBR and TLA) and the last one as a Stackelberg game (Local Markets). In this section, we introduce the basic concepts that will be used to derive the schemes.

5.3.1 Generalized Nash Equilibrium

Nash equilibrium problems (NEPs) [Nas51] arise in non-cooperative simultaneous games involving a set of players: in NEPs, the decisions of a player only affect the utility function of the other players. This notion is extended if, in addition, players have coupling constraints and the set of decisions of a player depends on the other players' decisions: this extension is referred to Generalized Nash Equilibrium Problems (GNEPs) [AD54]. In this work, we employ GNEPs in order to model TSO-DSO coordination.

5.3.1.1 Formulation of a GNEP

A Generalized Nash Equilibrium problem (GNEP) consists of a game among N players. Without loss of generality and for the sake of simplifying the exposition, we will only consider the case where $N = 2$. Player i controls variables $y_i,\ i = 1, 2$. y is the vector of all the variables: $y := (y_1,\ y_2)^\top$. The cost function of player i is denoted as π_i and can depend on the decisions of other players.

To optimize its strategy, each player minimizes its costs assuming that the strategy of the other player is fixed. This can be stated by the two following mathematical programs:
For $i \in \{1, 2\}, j \in \{1, 2\}, i \neq j$,

$$\mathcal{S}_i(y_j) : \min_{y_i} \quad \pi_i(y_i, y_j) \tag{5.2a}$$

$$s.t. \quad Y_i(y_i) \geq 0 \tag{5.2b}$$

$$(\beta_i) \quad B_i(y_i) + B_j(y_j) \geq 0 \tag{5.2c}$$

$\mathcal{S}_i(y_j)$ is the set of optimal solutions of the problem of player i depending on the decisions of player j. In (5.2a), each player is minimizing its cost. Each player has to obey individual private constraints (5.2b). Constraint (5.2c) is a global coupling constraint. β_i is the dual value associated to constraint (5.2c) for player $i = 1, 2$. We assume that (5.2) is convex which leads to a jointly convex GNEP. A solution of the GNEP, called generalized Nash equilibrium (or simply equilibrium), is a vector $y^* := (y_1^*,\ y_2^*)^\top$ such that $y_1^* \in \mathcal{S}_1(y_2^*)$ and $y_2^* \in \mathcal{S}_2(y_1^*)$. A specific equilibrium, namely a variational equilibrium (VE), is such that $\beta_1 = \beta_2$. Note that a VE does not always exist but the parallel between Variational Inequalities and GNEPs forms the basis of most solution concepts we introduce in the next section.

Remark. We could also consider that a player might not share truthfully its information with the other player. When dealing with asymmetric information, two configurations might arise:

(i) A player might have constraints involving their decisions and the decisions of the other player. These constraints are coupling but unknown for the other player. It naturally introduces *private coupling constraints* in the GNEP.

(ii) The players manipulate their report of information (the parameters of the coupling constraints). This is regarded as *strategic sharing of information*.

We do not explore asymmetric sharing of information in this chapter. Even if the assumption can be limiting in practice, the analysis of the schemes under this assumption is already of interest, especially in a GNEP setting. An

extension of this work considering a biased share of information (related to strategic sharing of information) between the TSO and the DSO is explored in [LCMP19].

5.3.1.2 Solution concepts

Proving existence, uniqueness or convergence of an algorithm to an equilibrium is a challenging task in the context of GNEP [FK10a]. Several methods have been proposed to solve GNEPs and their efficacy often depends on a special structure of the problem. Since GNEPs are naturally defined in a decentralized fashion, we first introduce decentralized methods. The most practical approaches rely on Jacobi iterative algorithms [PSFW08] or Gauss-Seidel-type algorithms [FK10a]. The main concept of these approaches relies on (i) considering the decisions of one player as fixed in order to deduce the decisions of the other and (ii) repeat this process until converging to an equilibrium. Classical optimization algorithms relying on similar concepts have recently been adapted to the case of GNEP such as penalty methods [FK10b] or augmented Lagrangian methods [KS16] and are particularly useful when focusing on VEs. Decentralized frameworks might be limiting if one aims at spanning a set of equilibria which is the case in our work.

Other methods consider a centralized paradigm. Most of them are based on the Nikaido-Isoda function [NI$^+$55] and Variational Inequalities. We decide to use one of them based on the theory developed by Nabetani, Tseng and Fukushima [NTF11] (we abbreviate this as NTF in the rest of the chapter). We motivate our choice by (i) the intuitiveness of the method, (ii) its direct application in our setting: a solution of the NTF problem is a GNE since we only consider global coupling equality constraints (the details are provided next, section 5.3.1.3), (iii) the fact that it is possible to span several equilibria for a particular simultaneous non-cooperative game.

The reader is referred to [FK10a] for more details about centralized and decentralized solution concepts for GNEPs.

5.3.1.3 The NTF method

Since the scope of the chapter focuses on agents coordinating to minimize their costs, there is typically a way to obtain a central socially optimal solution. Based on the formulation (5.2), the social cost (SC) is obtained by solving the

following constrained optimization problem:

$$\min_{y_1,y_2} \quad \pi_1(y_1,y_2)+\pi_2(y_1,y_2) \tag{5.3a}$$
$$s.t. \quad Y_1(y_1)\geq 0 \tag{5.3b}$$
$$Y_2(y_2)\geq 0 \tag{5.3c}$$
$$B_1(y_1)+B_2(y_2)\geq 0 \tag{5.3d}$$

If a solution of (5.3) is an equilibrium for the game setting chosen, a solution of (5.3) is a VE since both players have the same valuation of constraint (5.3d).

This special type of GNE is characterized by its link with Variational Inequalities [Ros65, Har91, FK10a, NTF11, LCJWA20]. VEs are not the only GNEs that can be computed. But solving a closely similar problem to (5.3) could be of interest to span equilibria. One way of computing variational equilibria and equilibrium candidates to the GNEP formulated in (5.2) is proposed by Nabetani, Tseng and Fukushima [NTF11] . The method relies on introducing parameters (γ_1,γ_2) and solving the following standard constrained optimization problem:

$$\mathcal{S}^{\mathrm{NTF}}(\gamma_1,\gamma_2): \min_{y_1,y_2} \quad \pi_1(y_1,y_2)+\pi_2(y_1,y_2)+\gamma_1 B_1(y_1)+\gamma_2 B_2(y_2) \tag{5.4a}$$
$$s.t. \quad Y_1(y_1)\geq 0 \tag{5.4b}$$
$$Y_2(y_2)\geq 0 \tag{5.4c}$$
$$B_1(y_1)+B_2(y_2)\geq 0 \tag{5.4d}$$

Finding a solution y^* of (5.4) is equivalent to finding y^* such that:

$$\langle \Pi(y^*,\gamma), y^*-y\rangle \geq 0, \quad \forall y\in\mathcal{F} \tag{5.5}$$

where:

$$\Pi(y,\gamma)=\nabla_y(\pi_1(y_1,y_2)+\pi_2(y_1,y_2)+\gamma_1 B_1(y_1)+\gamma_2 B_2(y_2))$$
$$\mathcal{F}=\{y \text{ satisfying (5.4b)} - \text{(5.4d)}\}$$

Problem (5.5) is a variational inequality problem which is parametrized by γ. We now explain the link between the variational inequality problem and the GNEP. To understand the intuition behind the use of the NTF to solve the GNEP, let us first focus on the case where $\gamma_1=\gamma_2=0$. In this case, (5.4) is (5.3), for which $\beta_1=\beta_2$. Assigning the same price is not necessary in general. Indeed, players 1 and 2 might assign different dual values to constraint (5.4d) depending on their own valuation of the global coupling constraint. By solving (5.4) for different values of γ_1 and γ_2, we are exploring potential solutions with different valuations of the coupling constraint by player 1 and player 2. This difference in valuation can also be interpreted as the influence of the player's

market power. This is why the NTF method is referred to 'a price-directed parametrized variational inequality' in [NTF11].

In [NTF11], the authors prove that:

$$\mathcal{S}^{\text{GNEP}} \subseteq \bigcup_{(\gamma_1,\gamma_2)\in\mathcal{W}} \mathcal{S}^{\text{NTF}}(\gamma_1,\gamma_2)$$

where $\mathcal{W} = \{(\gamma_1,\gamma_2) \in \mathbb{R}^2_+ \mid \gamma_1 \times \gamma_2 = 0\}$. In other words, it is possible to span the whole set of equilibria by testing all the possible values of γ_1 and γ_2 in $\mathcal{W}$. One should still check that the solutions obtained from problem (5.4) are actually equilibria (based on the definition of an equilibrium). Note also that from Theorem 3.3 in [NTF11], equality holds between the set of equilibria and the set of unions of NTF solutions if the shared constraints (5.4d) are equality constraints. Since we only consider global coupling equality constraints in the schemes of this work, the set of NTF solutions and GNEs are equal [NTF11]. No additional check is needed after solving (5.4) in our specific setting and no additional method is necessary to compute other equilibria.

In sections 5.4.1 and 5.4.2, we will proceed in casting the Shared Balancing Responsibility scheme and the TSO Limited Access scheme as GNEPs.

5.3.2 Stackelberg Games

Stackelberg games allow for modeling non-cooperative games with hierarchy in the agents' decision processes. Indeed, a Stackelberg game involves one or several leaders and one or several followers. We assume one leader and one follower in our case. In this configuration, the leader acts first before the follower reacts rationally to the signal sent by the leader. This hierarchical structure leads to modeling Stackelberg games as Bilevel Problems (BPs) [DKPVK15, DD12]. Denoting by y_L (resp. y_F) the set of decisions of the leader (resp. follower), the BP associated with a Stackelberg game is written as follows:

$$\begin{aligned}
&\min_{y_L, y_F(y_L)} && \pi_L(y_L) && && \text{(5.6a)}\\
&s.t. && Y_L(y_L) \geq 0 && && \text{(5.6b)}\\
& && B_L(y_L, y_F(y_L)) \geq 0 && && \text{(5.6c)}\\
& && y_F(y_L) = \quad \arg\min_{y_F} && \pi_F(y_F) && \\
& && \qquad\qquad s.t. && Y_F(y_F) \geq 0 && \text{(5.6d)}\\
& && && B_F(y_L, y_F) \geq 0 &&
\end{aligned}$$

Problem (5.6) shows how the lower level problem is embedded in the upper level problem. The lower level problem is parametrized by the upper level decision

y_L. The *rational reaction* function is usually introduced and is defined as:

$$\begin{aligned}\Phi(y_L) = \quad & \min_{y_F} \quad \pi_F(y_F) \\ & s.t. \quad Y_F(y_F) \geq 0 \\ & \qquad B_F(y_L, y_F) \geq 0\end{aligned}$$

Based on the use of the rational reaction function Φ, solving (5.6) is equivalent to solving the following problem:

$$\min_{y_L, y_F(y_L)} \quad \pi_L(y_L) \tag{5.7a}$$

$$s.t. \quad Y_L(y_L) \geq 0 \tag{5.7b}$$

$$B_L(y_L, y_F(y_L)) \geq 0 \tag{5.7c}$$

$$\Phi(y_L) \geq \pi_F(y_F) \tag{5.7d}$$

In general $\Phi()$ is non smooth which makes solving BPs challenging. As it is suggested in [DD12], when solving BPs, we usually rely on solving the mathematical programming with complementarity constraints (MPCC, also called mathematical programming with equilibrium constraints) problem associated with (5.6). The MPCC reformulation of (5.6) consists of writing the KKT conditions of the lower level problem at the upper level level and solve the upper level problem with (i) upper level primal variables and constraints (ii) primal and dual variables of the lower level problem (iii) KKT conditions of the lower level. MPCCs can be tackled by solvers such as Knitro [BNW06] or using the algorithm proposed in [YAO08]. Equivalence between BP and its MPCC reformulation is only true if the lower level problem satisfies Slater's condition for every possible value of y_L [DD12]. Since this condition is not trivial to verify in general, there is no guarantee that solving the MPCC provides solutions of (5.6). However, it is common practice to solve the MPCC and study if these solutions are equilibria for the original Stackelberg game.

We model the Local Markets (LM) scheme in section 5.4.3 as a Stackelberg game. We will illustrate the rational reaction approach and the MPCC approach in our numerical example.

5.3.3 Efficiency Loss and Price of Anarchy

To assess the quality and efficiency of the equilibria computed in section 5.5, we introduce the notion of Efficiency Loss (EL) defined for an equilibrium solution $\hat{y}$ as follows:

$$\text{EL}(\hat{y}) = \frac{\pi(\hat{y})}{SC^*}$$

where $\pi(\hat{y}) = \pi_1(\hat{y}_1) + \pi_1(\hat{y}_2)$ (or $\pi(\hat{y}) = \pi_L(\hat{y}_L) + \pi_F(\hat{y}_F)$) is the overall cost of equilibrium $\hat{y}$ and SC^* is the cost of the socially optimal solution (in the

scope of GNE and NTF introduced previously, it would be the cost of problem $\mathcal{S}^{NTF}(0,0)$ defined in (5.4)).

The notion of EL is naturally extended to evaluate a scheme by introducing the Price of Anarchy (PoA). Assuming that a certain coordination scheme CS modeled as a non-cooperative game leads to a set of equilibria Y^{CS}, the Price of Anarchy is defined as follows:

$$\text{PoA}(CS) = \max_{y \in Y^{CS}} EL(y)$$

In other words, the PoA is the 'worst' EL value taken by an equilibrium of a certain coordination scheme CS. The PoA takes value in $[1, +\infty[$ and a centralized efficient scheme is such that PoA $= 1$. Since the coordination schemes considered are decentralized, the PoA provides a measure of the efficiency loss incurred by the choice of a particular decentralized scheme. In practice, spanning the whole set of equilibria Y^{CS} or identifying the 'worst' equilibrium is difficult. In the numerical experiments of section 5.5, we report the PoA by identifying the worst equilibrium among the ones that we compute.

These measures will allow us to compare the equilibria and coordination schemes that we consider in this chapter.

5.4 Mathematical Formulations of the Coordination Schemes

In this section, we provide the game formulation of each scheme based on the centralized formulation of the T&D problem (5.2) and we rely on the game theory concepts detailed in section 5.3.

5.4.1 Shared Balancing Responsibilities

Each operator decouples the operations of its own network from the operations of the networks of other operators. We formulate this problem as a GNEP. As mentioned previously, the TSO is interacting with the DSO in a simultaneous game.

Considering $x^T = (p_{\mathcal{TG}}, f^{p,T}, \theta)$ as the decisions of the TSO and $x^D = (p_{\mathcal{DG}}, q, f^{p,D}, f^q, c, s)$ as the decisions of the DSO, since the interface constraint (5.1j) will be shared by both players, each problem is parametric on the decisions of the other problem; we then introduce SBR^{TSO} (resp. SBR^{DSO}) which represents the optimal value of the TSO (resp. DSO) problem depending on the DSO (resp. TSO) decisions. The associated problems are the following:

$$\text{SBR}^{\text{TSO}}(x^D) := \min_{x^T} \sum_{g \in \mathcal{TG}} C_g(p_g) \tag{5.8a}$$

$$s.t.\ \text{(5.1b)} - \text{(5.1d)}, \text{(5.1j)}, \text{(5.1k)} \tag{5.8b}$$

$$\mathrm{SBR}^{\mathrm{DSO}}(x^T) := \min_{x^D} \sum_{g \in \mathcal{DG}} C_g(p_g) \tag{5.9a}$$

$$s.t.\ \ (5.1\mathrm{e}) - (5.1\mathrm{i}), (5.1\mathrm{j}), (5.1\mathrm{l}) \tag{5.9b}$$

Remark. We show that a VE is an equilibrium for SBR. The details of the proof can be found in [MPLC18].

5.4.2 TSO Has Limited Access to DSO Resources

In this scheme, we allow the TSO to activate resources in the DSO's network. These resources bid in a transmission-level reserve activation market, in which the TSO does not account for DSO's network constraints. The TSO can then activate a certain amount of real power at distributed generator $g \in \mathcal{DG}$, through the decision variable $p_g^T,\ g \in \mathcal{DG}$. The TSO has a gaming opportunity in the sense that the TSO may choose to activate resources from the distribution network if it minimizes the overall transmission cost.

The DSO can also activate distributed resources, through the variable $p_g^D,\ g \in \mathcal{DG}$. The DSO is then responsible for operating the distribution network and takes into account transmission and distribution activation decisions in the distribution network. Nevertheless, since losses are non negligible in the distribution network, the DSO has to quantify how much power is flowing to the transmission network. Figure 5.1 illustrates this configuration. To that end, we introduce the variable $\eta_{i'j'}$. Thus, the variable $\eta_{i'j'}$ can be seen as a state variable capturing the losses incurred when the TSO is activating a certain amount of real power in the distribution network.

In this scheme, we then introduce the following constraints:

$$f_{i'j'}^{p,T} = \sum_{g \in \mathcal{DG}} p_g^T + \eta_{i'j'} \tag{5.10a}$$

$$f_{i'j'}^{p,D} = \sum_{g \in \mathcal{DG}} p_g^T + \eta_{i'j'} \tag{5.10b}$$

$$p_g = p_g^T + p_g^D, \qquad \forall g \in \mathcal{DG} \tag{5.10c}$$

Constraint (5.10a) (resp. (5.10b)) appears in the TSO (resp. DSO) problem. Variable $\eta_{i'j'}$ models the losses faced by the TSO when activating a certain amount of real power in the distribution network. Variable $\eta_{i'j'}$ is a state variable of the DSO which depends on the physics of the distribution network.

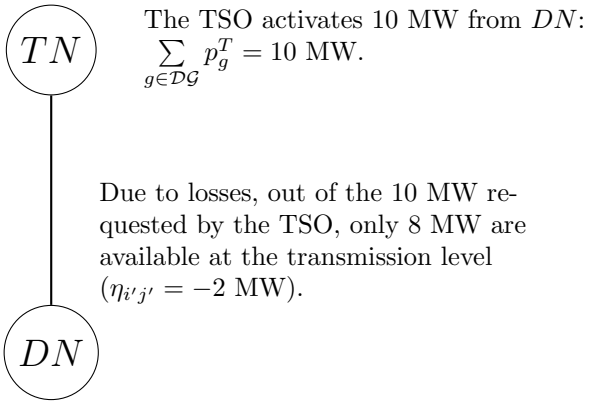


Figure 5.1: Illustration of how $\eta_{i'j'}$ is evaluated.

Constraint (5.10c) defines the generator injection in the distribution network as the sum of the decisions of the TSO and the DSO.

The two problems defining the GNEP are then:

$$\text{TLA}^{\text{TSO}}(y^{\text{DSO}}) := \min_{y^{TSO}=(x^T,p^T)} \sum_{g \in \mathcal{TG}} C_g(p_g) + \sum_{g \in \mathcal{DG}} C_g(p_g^T) \tag{5.11a}$$

$$s.t. \quad (5.1\text{b}) - (5.1\text{d}), (5.1\text{j}), (5.1\text{k}), (5.10\text{a}) \tag{5.11b}$$

$$\text{TLA}^{\text{DSO}}(y^{\text{TSO}}) := \min_{y^{DSO}=(x^D,p^D,\eta_{i'j'})} \sum_{g \in \mathcal{DG}} C_g(p_g^D) \tag{5.12a}$$

$$s.t. \quad (5.1\text{e}) - (5.1\text{i}), (5.1\text{j}), (5.1\text{l}), (5.10\text{b}), (5.10\text{c}) \tag{5.12b}$$

In particular, note that the TSO has the possibility to activate distribution resources but the DSO is responsible for making sure that it respects generation limits (equation (5.1i) includes constraints such as generation capacity constraints $\underline{p}_g \leq p_g = p_g^T + p_g^D \leq \overline{p}_g$).

5.4.3 Local Markets

In this coordination scheme, we assume that there is a separate local market, operated by the DSO. Resources from the DSO grid can only be offered to

the TSO after the DSO has selected resources needed to solve local imbalances within their periphery. The TSO is responsible for the operation of its own balancing market, where both resources from the transmission grid and resources from the distribution grid can participate. In practice, it can be a way to avoid the waste of power into the distribution grid, as well as helping in congestion management, and can be coupled with flexibility mechanisms like demand response. Contrary to the SBR and TLA introduced previously that we interpret as a non-cooperative game (implying that they both play at the same time), under the local markets coordination scheme that we describe in this section, the DSO is assumed to play first anticipating the rational reaction of the TSO, which reacts secondly to the signal sent by the DSO. We are then in a position of considering this scheme as a Stackelberg game where the DSO acts as the leader and the TSO is the follower, which means that the DSO clears its distribution market before the TSO clears the transmission market, taking into account transmission and distribution generation / consumption. With several distribution networks connected to the transmission network, the notion of Stackelberg equilibrium naturally extends to a multi-leader and single follower format. From section 5.3.2, we model this Stackelberg game as a BP as follows:

$$\min_{(x^T(x^D),x^D)} \quad \sum_{g\in\mathcal{DG}} C_g(p_g) \tag{5.13a}$$

$$s.t. \quad (5.1e)-(5.1i),(5.1j),(5.1l) \tag{5.13b}$$

$$x^T(x^D) = \quad \arg\min_{x^T} \quad \sum_{g\in\mathcal{TG}} C_g(p_g) \quad s.t. \quad (5.1b)-(5.1d),(5.1j),(5.1k) \tag{5.13c}$$

The lower level problem (5.13c) depends on the duplicate interface flow decision of the distribution $f^{p,D}_{i'j'}$.

5.5 Numerical Illustrations

5.5.1 A 2-bus Example

We consider the 2-bus example in Figure 5.2. We only consider real power flow here. The transmission bus gathers a 150 MW load and a generator g_1 with a capacity of 200 MW and a marginal cost of 20€. The line connecting the transmission and distribution buses has a capacity of 80 MW. The distribution bus consists of a 100 MW load and 2 generators: g_2, with a capacity of 50 MW and a marginal cost of 10€, and g_3, with a 100 MW capacity and a 25€ marginal cost. Distribution generation is usually cheaper than transmission generation, but losses might sometimes significantly increase the cost of generation in the

distribution and that is why we consider one cheap generator and another one more expensive.

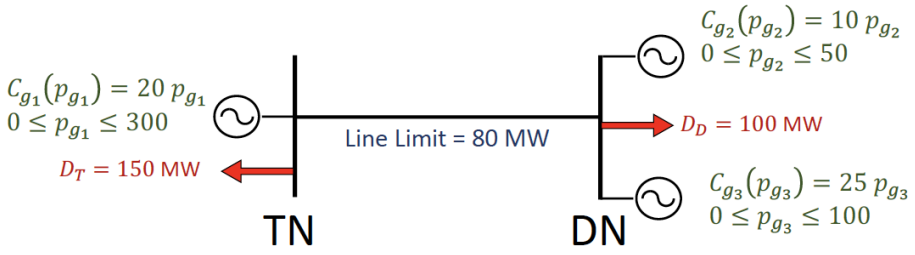


Figure 5.2: The toy example considered for testing the different coordination schemes.

The centralized optimization problem associated with the 2-bus example is written as follows:

$$\min\ 20p_{g_1} + 10p_{g_2} + 25p_{g_3} \tag{5.14a}$$
$$s.t.\quad p_{g_1} - 150 = f^T \tag{5.14b}$$
$$0 \le p_{g_1} \le 300 \tag{5.14c}$$
$$p_{g_2} + p_{g_3} - 100 = -f^D \tag{5.14d}$$
$$0 \le p_{g_2} \le 50 \tag{5.14e}$$
$$0 \le p_{g_3} \le 100 \tag{5.14f}$$
$$f^T = f^D \tag{5.14g}$$
$$-80 \le f^T \le 80 \tag{5.14h}$$
$$-80 \le f^D \le 80 \tag{5.14i}$$

The equilibria obtained are displayed in Figure 5.3 (as a function of the operators' costs) and Figure 5.4 (as a function of the generation decisions). Note that some points are equilibria for one or several schemes as shown in the legend of Figures 5.3 and 5.4. Note also that we report the EL of each equilibrium in Figure 5.3. We discuss the computation of the equilibria and their efficiency next.

5.5.1.1 Centralized Solution

The centralized solution of the 2-bus example clears the market as follows: the cheapest distribution generator produces at its maximum capacity ($p_{g_2} = 50$) and the rest is left to the transmission generator ($p_{g_1} = 200$) to compensate the transmission imbalance of 150 MW and 50 MW are flowing at the interconnection to the distribution network to satisfy the 100 MW distribution load. This

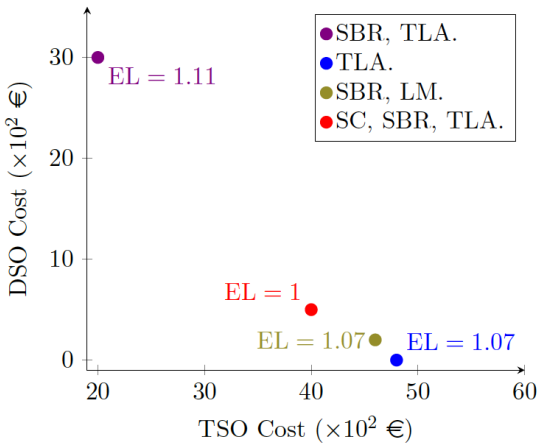


Figure 5.3: Equilibria spanned for the 2-bus example as a function of the operators' costs. We also report the EL of each equilirbium.

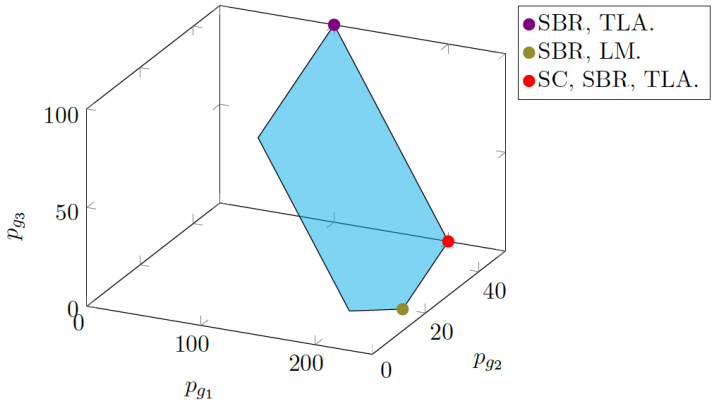


Figure 5.4: Equilibria spanned for the 2-bus example as a function of the generation decisions. The cyan area shows the feasibility set of the problem.

dispatch leads to a total cost of 4500€ and is represented in red in Figures 5.3 and 5.4 as the minimizing social cost (SC) solution. Note that the SC solution is unique and is the only VE for this example.

5.5.1.2 SBR Equilibria

We write the NTF optimization problem associated with the SBR problems (5.8) and (5.9) as:

$$\begin{aligned} \mathrm{NTF}_{\mathrm{SBR}}(\gamma_T, \gamma_D) : \min \quad & 20p_{g_1} + 10p_{g_2} + 25p_{g_3} + \gamma_T f^T - \gamma_D f^D \\ s.t. \quad & (5.14\mathrm{b}) - (5.14\mathrm{i}) \end{aligned} \tag{5.15}$$

From [NTF11], it is sufficient to study values in $\mathcal{W} = \{(\gamma_T, \gamma_D) \in \mathbb{R}^2_+ \mid \gamma_T \gamma_D = 0\}$. We then have to look into three different cases:

1. $\underline{\gamma_T = \gamma_D = 0}$. This case is straightforward and (5.15) is equivalent to (5.14) represented in red in Figures 5.3 and 5.4.
2. $\underline{\gamma_T > 0,\ \gamma_D = 0}$. In this case, one interpretation of the objective function is that we are implicitly 'penalizing the exports from the transmission network'. When substituting f^T in the objective by its expression in (5.14b), the marginal cost of p_{g_1} becomes $20 + \gamma_T$. If this marginal cost is lower than the marginal cost of p_{g_3}, the deduced equilibrium is SC (red). This condition is met when $\gamma_T \leq 5$. If $\gamma_T > 5$, the transmission exports become more expensive than generator g_3 and the equilibrium solution is $p_{g_1} = 100,\ p_{g_2} = 50,\ p_{g_3} = 100$ (violet point).
3. $\underline{\gamma_T = 0,\ \gamma_D > 0}$. Now we are 'penalizing the distribution exports'. The same reasoning can be conducted and we deduce the SC equilibrium if $\gamma_D \leq 10$. If $\gamma_D > 10$, we deduce another equilibrium, where one aims at minimizing distribution exports $p_{g_1} = 230,\ p_{g_2} = 20,\ p_{g_3} = 0$ (green point).

The interpretation of the SBR scheme is fairly straightforward. Depending on the valuation of the interface flow, if this flow does not correspond to the optimal flow it will result in sub-optimal dispatch. The specific choice of flow at the interface will determine the extent of efficiency losses of the SBR scheme.

5.5.1.3 TLA Equilibria

We write the NTF optimization problem associated with the TLA problems (5.11) and (5.12) as:

$$\begin{aligned} \mathrm{NTF}_{\mathrm{TLA}}(\gamma_T, \gamma_D) : \min \quad & 20p_{g_1} + 10p_{g_2} + 25p_{g_3} + \gamma_T f^T - \gamma_D f^D \\ s.t. \quad & (5.14\mathrm{b}) - (5.14\mathrm{i}) \\ & p_{g_2} = p^T_{g_2} + p^D_{g_2} \\ & p_{g_3} = p^T_{g_3} + p^D_{g_3} \\ & f^T = p^T_{g_2} + p^T_{g_3} \\ & f^D = p^T_{g_2} + p^T_{g_3} \end{aligned} \tag{5.16}$$

Using the same reasoning as SBR, we have to look into three different cases:

1. $\gamma_T = \gamma_D = 0$. This case is straightforward and (5.16) is equivalent to (5.14) represented in red in Figures 5.3 and 5.4.

2. $\gamma_T > 0,\ \gamma_D = 0$. In this case, we are implicitly penalizing the exports from the transmission network as well as the use of distribution resources by the TSO. Again, if $\gamma_T > 5$, the transmission exports become more expensive than generator g_3 and the equilibrium solution is $p_{g_1} = 100,\ p^D_{g_2} = 50,\ p^T_{g_2} = 0,\ p^D_{g_3} = 100,\ p^T_{g_3} = 0$ (violet point).

3. $\gamma_T = 0,\ \gamma_D > 0$. Now we are penalizing the distribution exports but also encouraging transmission activations in the distribution network. The same reasoning can be conducted and we deduce the SC equilibrium if $\gamma_D \leq 10$. If $\gamma_D > 10$, we deduce a different equilibrium, where the TSO covers all the costs of activations $p_{g_1} = 230,\ p^D_{g_2} = 0,\ p^T_{g_2} = 20,\ p^D_{g_3} = 0,\ p^T_{g_3} = 0$ (blue point[1]).

5.5.1.4 LM Equilibrium

First, let us write the TSO's lower level problem, which is parametric of the decisions of the DSO, and the KKT conditions of the problem:

$$
\begin{array}{lll}
\min & 20 p_{g_1} & \\
s.t. & p_{g_1} - 150 = f^T & (\lambda) \\
& 0 \leq p_{g_1} \leq 300 & (\rho^-, \rho^+) \\
& f^T = f^D & (\mu) \\
& -80 \leq f^T \leq 80 & (\phi^-, \phi^+)
\end{array}
\qquad
\begin{array}{l}
20 + \lambda - \rho^- + \rho^+ = 0 \\
-\lambda - \phi^- + \phi^+ = 0 \\
p_{g_1} - 150 = f^T \\
f^T = f^D \\
p_{g_1} \geq 0 \perp \rho^- \geq 0 \\
p_{g_1} \leq 300 \perp \rho^+ \geq 0 \\
f^T \geq -80 \perp \phi^- \geq 0 \\
f^T \leq 80 \perp \phi^+ \geq 0
\end{array}
$$

In this case, we can easily deduce a rational reaction function, and the MPCC would also result in the problem to solve. We can express TSO's decisions as a function of the DSO's decisions as follows:

$$
\begin{aligned}
f^T &= f^D \\
p_{g_1} &= 150 + f^D
\end{aligned}
$$

[1]This point is not displayed on Figure 5.4 because of the duplication of the variable p in the distribution network. However, in terms of pure activation, the blue point would be the same one as the green point on Figure 5.4

The upper level is then written as:

$$\begin{aligned} \min \quad & 10p_{g_2} + 25p_{g_3} \\ s.t. \quad & p_{g_2} + p_{g_3} - 100 = -f^D \\ & 0 \leq p_{g_2} \leq 50 \\ & 0 \leq p_{g_3} \leq 100 \\ & -80 \leq f^D \leq 80 \\ & f^T = f^D \\ & p_{g_1} = 150 + f^D \end{aligned}$$

The solution of this Stackelberg game is unique and displayed in green in Figure 5.3. The solution is slightly suboptimal in this case (total activation cost: 4800€, EL = 1.07). Indeed, the transmission bus is activating more power (p_{g_1} = 230 MW) to compensate the lighter activation from generator 2 (p_{g_2} = 20 MW). These decisions lead to congesting the interface line. This solution can be interpreted as the fact the DSO is anticipating a higher penetration of transmission generation in the distribution network. By doing so, the TSO is reacting to compensate the imbalance in the distribution network. Even if appealing from the theoretical point of view, the scheme does not capture the impact of the transmission resources. This conclusion is also drawn and further explained in [LCMP19]. The original goal being to fully use RES in the distribution network, one idea could have been to penalize the interconnection flow in the DSO problem in order to fully activate resources in the distribution network. However, this would lead to activating the expensive generation unit in the distribution network.

Remark. Another idea would have been to consider errors in the rational reaction of the TSO. In a way, it could capture the fact that the TSO might not react rationally to the signal sent by the DSO. We do not further explore '*bounded rationality*' in this chapter and decide to stay in the classical scope of game theory with profit maximizing rational agents. However, preliminary study in a bounded rational setting is conducted in [LCMP19].

Remark. One outcome of the results of this scheme is that one way to make a sequential framework profitable for the TSO and the DSO could be to appropriately valuate the interface flow. It is what has been explored in chapter 4.

This preliminary study shows that each scheme can generate some inefficiencies, with LM exhibiting superior performance than the other schemes because it provides a unique equilibrium (PoA(LM) = 1.07) which is slightly inefficient, but more efficient than the worst equilibria of SBR and TLA (PoA(SBR) =

PoA(TLA) = 1.07). In order to further elaborate on the comparison of TLA and SBR, we aim at extending these results on a more representative example.

5.5.2 Extending the 2-bus Results

We also show equilibria of SBR and TLA on the test case represented in Figure 3.2. We expand equilibria using the NTF method. The results are shown in Figures 5.5 and 5.6. Again, the SC solution is unique and is the only VE. We span as many equilibria as possible to identify all the potential outcomes of each scheme.

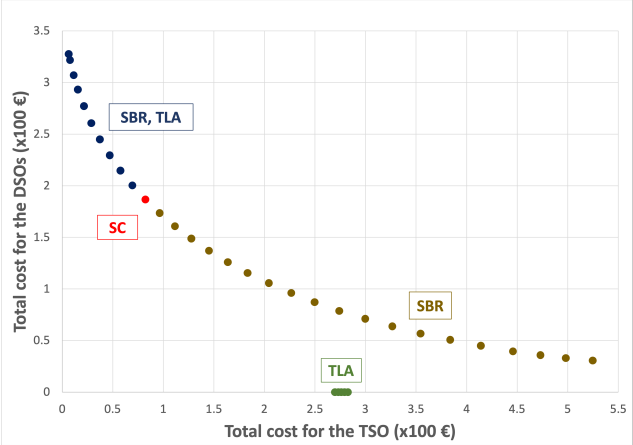


Figure 5.5: Equilibria for SBR and TLA as a function of the operators' cost.

Figure 5.5 generalizes Figure 5.2: indeed, except for the SC solution, the points drawn in Figure 5.2 expand as a group of points where the equilibria have different interpretations that are discussed next.

Figure 5.6 ranks the EL of each equilibrium obtained. It is interesting to see that the EL ranges from 1 to 2.1, which demonstrates the potential inefficiency of some equilibria, especially for SBR.

5.5.2.1 SBR Equilibria

The interpretation of the SBR scheme is fairly straightforward. Once the flow of real power is fixed at the T&D interface, if this flow does not correspond to the optimal flow it will result in a sub-optimal dispatch. By fixing the interface flow, the SBR scheme decouples the problems of the agents and therefore results in an equilibrium, unless the value of the interface flow cannot be supported by a physically feasible dispatch at any of the sub-networks. The specific choice

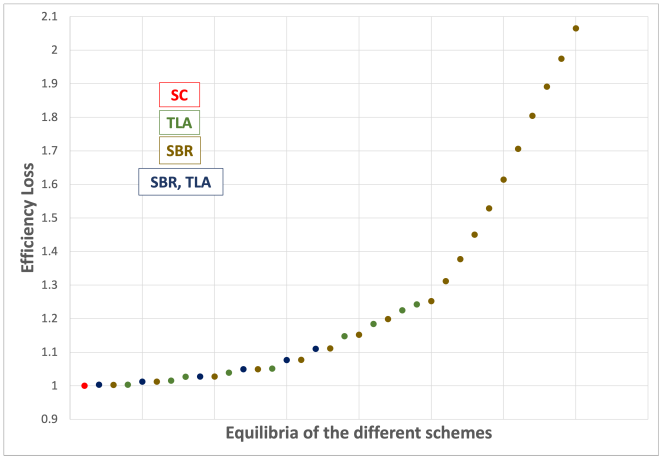


Figure 5.6: EL of the equilibria spanned for SBR and TLA. The equilibria are ranked from smallest to largest EL.

of flow at the interface will determine the extent of efficiency losses of the SBR scheme. Figure 5.6 shows how SBR can lead to inefficient equilibria, in particular, from Figure 5.6, PoA(SBR) = 2.07. The SBR scheme could eventually be improved if we were to consider a third player in charge of efficiently determining the power flow at the interface or through a intensive and careful use of historical data to settle the interface flow.

5.5.2.2 TLA Equilibria

For this coordination scheme, the equilibria to the left of the welfare solution are identical to the corresponding equilibria of SBR. This is the case in which the TSO does not activate distribution resources and let this responsibility to the DSO. The interpretation of the right-hand side cluster of equilibria is more delicate. These are equilibria for which the TSO finds it worthwhile to activate such a large quantity of distribution network reserves, that these reserves are sufficient for both covering the distribution network disturbances, while also supporting the transmission network disturbances, even if part of the power is foregone as resistive losses on the DSO network (with equilibrium points further to the right corresponding to increasing real power losses on the DSO network). This turns out to be an equilibrium because (i) clearly for the DSOs there is no reason to deviate, since the TSO is paying for reserves that cover the DSOs' imbalance, while (ii) even when covering the DSOs imbalance and facing distribution system losses, the TSO still finds it preferable to activate distributed resources due to the fact that the transmission system reserve activation is a

more expensive alternative.

5.6 Conclusion

This chapter is focused on the modeling of TSO-DSO coordination schemes as non-cooperative games. A centralized dispatch of the entire system is challenging due to the large size of the network in general. Decentralized models, where the computational effort would be separated and the privacy of information of each operator would be preserved, may be more viable in practice. We employ GNE in order to quantify the efficiency losses of two alternative TSO-DSO coordination schemes and analyze the results on a small-scale example as well as a third scheme that introduced sequentiality. We unveil a multiplicity of equilibria for each scheme and unveil a free-riding effect in the TLA coordination scheme, whereby for some equilibria the DSO imbalances are entirely covered by the TSO. LM provided a suboptimal dispatch because of some unwanted effect of the anticipation of the DSO because the DSO sees the interface flow as a free resource.

Conclusions

As mentioned in the introduction, five aspects of the reorganization of electricity markets are explored in this dissertation: (i) the uncertainty of RES, (ii) the complexity of distribution networks, (iii) the increasing importance of the role of the DSO, (iv) TSO-DSO coordination, and (v) the computation of consistent electricity prices.

In chapter 2, we focus on developing algorithms to solve the well-known optimal power flow problem. We exclusively considered the AC version of the problem (because it is necessary in distribution networks) and first looked at a deterministic version before proposing a framework for the stochastic variant. For the deterministic AC-OPF, we suggest a Gauss-Newton algorithm which relies on an ℓ_1-penalization of the complicating constraints. We provide theoretical guarantees for the algorithm to converge. We also explain the parametrization of the method and demonstrate comparable performances with Ipopt on a set of state-of-the-art instances (MATPOWER). The Gauss-Newton scheme could also improve in performance with more stable ADMM solvers. In the stochastic version of the problem, we assume that the demand is uncertain, which is more realistic in practice due to RES proliferation and flexibility mechanisms (like demand response). The suggested method, which solves a chance constrained form of AC-OPF, rapidly computes a solution that is enhanced iteratively by intelligently working with a reduced scenario set. The data-driven approach is validated on a 1354-bus system where only 30 critical scenarios were necessary to compute a sufficiently robust solution.

Chapter 3 exposes the preliminary content on TSO-DSO coordination. With a more involved DSO, we derive five coordination schemes which transpose different paradigms for a future market organization. The schemes incorporate one transmission network, with a DC-OPF model, and one or several radial distribution networks on which the second-order cone relaxation of OPF is applied. A simple numerical experiment allows us to already stress advantages and drawbacks of each of the schemes.

In chapter 4, we focus on the scheme 'Decentralized Common Market' in depth. This chapter aims at developing a practical platform in line with the future vision of how transmission and distribution should be operated in Eu-

rope. The platform, based on the idea of constructing Residual Supply Functions (RSFs), offers several compelling features: (i) the distribution networks are precisely modeled using AC power flow models, (ii) complex bid structure (such as block bids) is considered, (iii) the platform preserves operators' privacy by only sharing minimal border information (iv) dispatch decisions but also prices are returned by the envisioned clearing platform, and (v) the framework is largely parrallelizable. The platform is tested and validated on national test systems of Italy and Denmark, which scale to thousands of buses and tens of thousands of bids.

Other coordination schemes are further explored from a game theoretical point of view in chapter 5, namely 'Shared Balancing Responsibilities' and 'Local Ancillary Services Market'. The schemes are modeled as non-cooperative simultaneous or sequential games and the generalized Nash equilibria computed stress the efficiency of the considered schemes. Two illustrative examples demonstrate the different market situations in which the schemes might lead to.

Bibliography

[AD54] Kenneth J. Arrow and Gerard Debreu. Existence of an Equilibrium for a Competitive Economy. *Econometrica: Journal of the Econometric Society*, pages 265–290, 1954.

[ART03] Erling D. Andersen, Cornelis Roos, and Tamas Terlaky. On Implementing a Primal-Dual Interior-Point Method for Conic Quadratic Optimization. *Mathematical Programming*, 95(2):249–277, 2003.

[ASL+17] Araz Ashouri, Peter Sels, Guillaume Leclercq, Olivier Devolder, Frederik Geth, and Reinhilde D'Hulst. Market Design for Centralized Coordination Mechanism. Technical Report Deliverable 2.4, SmartNet, 2017.

[BBC+19] Sogol Babaeinejadsarookolaee, Adam Birchfield, Richard D. Christie, Carleton Coffrin, Christopher DeMarco, Ruisheng Diao, Michael Ferris, Stephane Fliscounakis, Scott Greene, Renke Huang, et al. The Power Grid Library for Benchmarking AC Optimal Power Flow Algorithms. *arXiv preprint arXiv:1908.02788*, 2019.

[BCH14] Daniel Bienstock, Michael Chertkov, and Sean Harnett. Chance-constrained Optimal Power Flow: Risk-aware Network Control Under Uncertainty. *Siam Review*, 56(3):461–495, 2014.

[BDL19] Digvijay Boob, Qi Deng, and Guanghui Lan. Proximal Point Methods for Optimization with Nonconvex Functional Constraints. *arXiv preprint arXiv:1908.02734*, 2019.

[BNW06] Richard H. Byrd, Jorge Nocedal, and Richard A. Waltz. Knitro: An Integrated Package for Nonlinear Optimization. In *Large-scale nonlinear optimization*, pages 35–59. Springer, 2006.

[BP16] Jérôme Bolte and Edouard Pauwels. Majorization-Minimization Procedures and Convergence of SQP Methods for Semi-Algebraic and Tame Programs. *Mathematics of Operations Research*, 41(2):442–465, 2016.

[BPC+11] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122, 2011.

[BVC20] Lucien Bobo, Andreas Venzke, and Spyros Chatzivasileiadis. Second-Order Cone Relaxations of the Optimal Power Flow for Active Distribution Grids. *arXiv preprint arXiv:2001.00898*, 2020.

[Car62] J. Carpentier. Contribution à l'étude du dispatching économique. *Bulletin de la Société Française des Electriciens*, 3(1):431–447, 1962.

[CBS+18a] Carleton Coffrin, Russell Bent, Kaarthik Sundar, Yeesian Ng, and Miles Lubin. PowerModels.jl: An Open-Source Framework for Exploring Power Flow Formulations. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–8. IEEE, 2018.

[CBS+18b] Carleton Coffrin, Russell Bent, Kaarthik Sundar, Yeesian Ng, and Miles Lubin. PowerModels.jl: An Open-Source Framework for Exploring Power Flow Formulations. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–8, June 2018. doi:10.23919/PSCC.2018.8442948.

[CC06] Giuseppe C. Calafiore and Marco C. Campi. The Scenario Approach to Robust Control Design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006.

[CDT+15] Mohammadreza Chamanbaz, Fabrizio Dabbene, Roberto Tempo, Venkatakrishnan Venkataramanan, and Qing-Guo Wang. Sequential Randomized Algorithms for Convex Optimization in the Presence of Uncertainty. *IEEE Transactions on Automatic Control*, 61(9):2565–2571, 2015.

[CGT14] Coralia Cartis, Nicholas I. M. Gould, and Philippe L. Toint. On the Complexity of Finding First-Order Critical Points in Constrained Nonlinear Optimization. *Mathematical Programming*, 144(1-2):93–106, 2014.

[CH11] Duncan Callaway and Ian Hiskens. Achieving Controllability of Electric Loads. *Proceedings of the IEEE*, 99(1):184–199, 2011.

[CHVH15] Carleton Coffrin, Hassan L. Hijazi, and Pascal Van Hentenryck. The QC Relaxation: A Theoretical and Computational Study on Optimal Power Flow. *IEEE Transactions on Power Systems*, 31(4):3008–3018, 2015.

[CNH+16] Michael Caramanis, Elli Ntakou, William W. Hogan, Aranya Chakrabortty, 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*, 104(4):807–836, 2016.

[Com17] The European Commission. Regulation of the European Parliament and of the Council: on the Internal Market for Electricity, 2017. URL: https://eur-lex.europa.eu/resource.html?uri=cellar:9b9d9035-fa9e-11e6-8a35-01aa75ed71a1.0012.02/DOC_1&format=PDF.

[Com19] The European Commission. Communication from the Commission: The European Green Deal, 2019. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1576150542719&uri=COM%3A2019%3A640%3AFIN.

[Com20a] The European Commission. ASSET Study on Regulatory Priorities for Enabling Demand Side Flexibility, 2020. URL: https://op.europa.eu/en/publication-detail/-/publication/38fbccf2-35de-11eb-b27b-01aa75ed71a1.

[Com20b] The European Commission. Report from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Renewable Energy Progress Report, 2020. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1602743359876&uri=COM:2020:952:FIN.

[CP11] Antonin Chambolle and Thomas Pock. A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.

[CRS18] Frank E. Curtis, Daniel P. Robinson, and Mohammadreza Samadi. Complexity Analysis of a Trust Funnel Algorithm for Equality Constrained Optimization. *SIAM Journal on Optimization*, 28(2):1533–1563, 2018.

[CS18] Antonio J. Conejo and Ramteen Sioshansi. Rethinking Restructured Electricity Market Design: Lessons Learned and Future

Needs. *International Journal of Electrical Power & Energy Systems*, 98:520–530, 2018.

[CSY19] Yifan Chen, Yuejiao Sun, and Wotao Yin. Run-and-Inspect Method for Nonconvex Optimization and Global Optimality Bounds for R-local Minimizers. *Mathematical Programming*, 176(1-2):39–67, 2019.

[DBC17] Deepjyoti Deka, Scott Backhaus, and Michael Chertkov. Structure Learning in Power Distribution Networks. *IEEE Transactions on Control of Network Systems*, 5(3):1061–1074, 2017.

[DBS17] Emiliano Dall'Anese, Kyri Baker, and Tyler Summers. Chance-constrained AC Optimal Power Flow for Distribution Systems with Renewables. *IEEE Transactions on Power Systems*, 32(5):3427–3438, 2017.

[DD12] Stephan Dempe and Joydeep Dutta. Is Bilevel Programming a Special Case of a Mathematical Program with Complementarity Constraints? *Mathematical Programming*, 131(1-2):37–48, 2012.

[DGKR03] Jitka Dupačová, Nicole Gröwe-Kuska, and Werner Römisch. Scenario Reduction in Stochastic Programming. *Mathematical Programming*, 95(3):493–511, 2003.

[DKPVK15] Stephan Dempe, Vyacheslav Kalashnikov, Gerardo A. Pérez-Valdés, and Nataliya Kalashnykova. Bilevel Programming Problems. *Energy Systems. Springer, Berlin*, 2015.

[DM19] D. Deka and S. Misra. Learning for DC-OPF: Classifying Active Sets Using Neural Nets. In *2019 IEEE Milan PowerTech*, pages 1–6, 2019.

[DOW+55] George B. Dantzig, Alex Orden, Philip Wolfe, et al. The Generalized Simplex Method for Minimizing a Linear Form Under Linear Inequality Restraints. *Pacific Journal of Mathematics*, 5(2):183–195, 1955.

[DP94] Gianni Di Pillo. Exact Penalty Methods. In *Algorithms for Continuous Optimization*, pages 209–253. Springer, 1994.

[DP19] Dmitriy Drusvyatskiy and Courtney Paquette. Efficiency of Minimizing Compositions of Convex Functions and Smooth Maps. *Mathematical Programming*, 178(1-2):503–558, 2019.

[DPG+14] Yann N. Dauphin, Razvan Pascanu, Caglar Gulcehre, Kyunghyun Cho, Surya Ganguli, and Yoshua Bengio. Identifying and Attacking the Saddle Point Problem in High-Dimensional Non-Convex Optimization. In *Advances in Neural Information Processing Systems*, pages 2933–2941, 2014.

[DPN+13] Yi Ding, Salvador Pineda, Preben Nyeng, Jacob Ostergaard, Emil Larsen, and Quiwei Wu. Real-Time Market Concept Architecture for EcoGrid EU - A Prototype for European Smart Grids. *IEEE Transactions on Smart Grid*, 4(4):2006–2016, 2013.

[DVSD+19] Kristof De Vos, Nicolas Stevens, Olivier Devolder, Anthony Papavasiliou, Bob Hebb, and James Matthys-Donnadieu. Dynamic Dimensioning Approach for Operating Reserves: Proof of Concept in Belgium. *Energy Policy*, 124:272–285, 2019.

[EDA18] Anders Eltved, Joachim Dahl, and Martin S. Andersen. On the Robustness and Scalability of Semidefinite Relaxation for Optimal Power Flow Problems. *Optimization and Engineering*, pages 1–18, 2018.

[EGG17] Calum Edmunds, Stuart Galloway, and Simon Gill. Distributed Electricity Markets and Distribution Locational Marginal Prices: A Review. In *2017 52nd International Universities Power Engineering Conference (UPEC)*, pages 1–6. IEEE, 2017.

[EN09] Andreas Ehrenmann and Karsten Neuhoff. A Comparison of Electricity Market Designs in Networks. *Operations research*, 57(2):274–286, 2009.

[ER99] Antonio Gómez Expósito and E. Romero Ramos. Reliable Load Flow Technique for Radial Distribution Networks. *IEEE Transactions on Power Systems*, 14(3):1063–1069, 1999.

[FK10a] Francisco Facchinei and Christian Kanzow. Generalized Nash Equilibrium Problems. *Annals of Operations Research*, 175(1):177–211, 2010.

[FK10b] Francisco Facchinei and Christian Kanzow. Penalty Methods for the Solution of Generalized Nash Equilibrium Problems. *SIAM Journal on Optimization*, 20(5):2228–2253, 2010.

[FKP+14] Hans Joachim Ferreau, Christian Kirches, Andreas Potschka, Hans Georg Bock, and Moritz Diehl. qpOASES: A Parametric Active-Set Algorithm for Quadratic Programming. *Mathematical Programming Computation*, 6(4):327–363, 2014.

[FL13] Masour Farivar and Steven H. Low. Branch Flow Model: Relaxations and Convexification - Part I. *IEEE Transactions on Power Systems*, 28(3):2554–2564, 2013.

[fSG14] European Distribution System Operators for Smart Grids. Flexibility: The Role of DSOs in Tomorrow's Electricity Market. Technical report, 2014.

[FSR12] Stephen Frank, Ingrida Steponavice, and Steffen Rebennack. Optimal Power Flow: A Bibliographic Survey I. *Energy Systems*, 3(3):221–258, 2012.

[GHP+07] Paul R. Gribik, William W. Hogan, Susan L. Pope, et al. Market-Clearing Electricity Prices and Energy Uplift. *Cambridge, MA*, 2007.

[Gil14] Hans Christian Gils. Assessment of the Theoretical Demand Response Potential in Europe. *Energy*, 67:1–18, 2014.

[GKHR03] Nicole Growe-Kuska, Holger Heitsch, and Werner Romisch. Scenario Reduction and Scenario Tree Construction for Power Management Problems. In *2003 IEEE Bologna Power Tech Conference Proceedings,*, volume 3, pages 7–pp. IEEE, 2003.

[GLTL14] Lingwen Gan, Na Li, Ufuk Topcu, and Steven H. Low. Exact Convex Relaxation of Optimal Power Flow in Radial Networks. *IEEE Transactions on Automatic Control*, 60(1):72–87, 2014.

[GNB20] Manuel Garcia, Harsha Nagarajan, and Ross Baldick. Generalized Convex Hull Pricing for the AC Optimal Power Flow Problem. *IEEE Transactions on Control of Network Systems*, 7(3):1500–1510, 2020.

[GO19] LLC Gurobi Optimization. Gurobi Optimizer Reference Manual, 2019. URL: `http://www.gurobi.com`.

[GPS18] Helena Gerard, Enrique Israel Rivero Puente, and Daan Six. Coordination Between Transmission and Distribution System Operators in the Electricity Sector: A Conceptual Framework. *Utilities Policy*, 50:40–48, 2018.

[GRS16] Helena Gerard, Enrique Rivero, and Daan Six. Basic Schemes for TSO-DSO Coordination and Ancillary Services Provision. Technical report, SmartNet, 2016.

[Har91] Patrick T. Harker. Generalized Nash Games and Quasi-Variational Inequalities. *European Journal of Operational research*, 54(1):81–94, 1991.

[HB16] Bowen Hua and Ross Baldick. A Convex Primal Formulation for Convex Hull Pricing. *IEEE Transactions on Power Systems*, 32(5):3814–3823, 2016.

[Hob01] Benjamin F. Hobbs. Linear complementarity Models of Nash-Cournot Competition in Bilateral and POOLCO Power Markets. *IEEE Transactions on power systems*, 16(2):194–202, 2001.

[HS18] Lion Hirth and Ingmar Schlecht. Market-Based Redispatch in Zonal Electricity Markets, 2018.

[HSZ+19] Bowen Hua, Dane A. Schiro, Tongxin Zheng, Ross Baldick, and Eugene Litvinov. Pricing in Multi-Interval Real-Time Markets. *IEEE Transactions on Power Systems*, 34(4):2696–2705, 2019.

[HWO+15] Shaojun Huang, Quiwei Wu, Shmuel S. Oren, Ruoyang Li, and Zhaoxi Lu. Distribution Locational Marginal Pricing through Quadratic Programming for Congestion Management in Distribution Networks. *IEEE Transactions on Power Systems*, 30(4):2170–2178, 2015.

[Jab06] Rabih A. Jabr. Radial Distribution Load Flow using Conic Programming. *IEEE Transactions on Power Systems*, 21(3):1458–1459, 2006.

[Jab07] Rabih A. Jabr. A Conic Quadratic Format for the Load Flow Equations of Meshed Networks. *IEEE Transactions on Power Systems*, 22(4):2285–2286, Nov 2007. doi:10.1109/TPWRS.2007.907590.

[Jab08] Rabih A. Jabr. Optimal Power Flow Using an Extended Conic Quadratic Formulation. *IEEE Transactions on Power Systems*, 23(3):1000–1008, 2008.

[JT07] Paul Joskow and Jean Tirole. Reliability and Competitive Electricity Markets. *The Rand Journal of Economics*, 38(1):60–84, 2007.

[JYS99] Wei Jing-Yuan and Yves Smeers. Spatial Oligopolistic Electricity Models with Cournot Generators and Regulated Transmission Prices. *Operations Research*, 47(1):102–112, 1999.

[KCL+14] Matt Kraning, Eric Chu, Javad Lavaei, Stephen P. Boyd, et al. *Dynamic Network Energy Management via Proximal Message Passing*. Now Publishers, 2014.

[KDS16] 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.

[Kel99] Carl T. Kelley. *Iterative Methods for Optimization*, volume 18. SIAM, 1999.

[KFS18] Drosos Kourounis, Alexander Fuchs, and Olaf Schenk. Toward the Next Generation of Multiperiod Optimal Power Flow Solvers. *IEEE Transactions on Power Systems*, 33(4):4005–4014, 2018.

[KMT16] Lorenzo Kristov, Paul De Martini, and Jeffrey Taft. A Tale of Two Visions: Designing a Decentralized Transactive Electric System. *IEEE Power and Energy Magazine*, 14(3):63–69, 2016.

[KS16] Christian Kanzow and Daniel Steck. Augmented Lagrangian Methods for the Solution of Generalized Nash Equilibrium Problems. *SIAM Journal on Optimization*, 26(4):2034–2058, 2016.

[KZ13] Friedrich Kunz and Alexander Zerrahn. The Benefit of Coordinating Congestion Management in Germany. In *European Energy Market (EEM), 2013 10th International Conference on the*, pages 1–8. IEEE, 2013.

[LA16] George Liberopoulos and Panagiotis Andrianesis. Critical Review of Pricing Schemes in Markets with Non-convex Costs. *Operations Research*, 64(1):17–31, 2016.

[LÁT19] Leandro Lind, José Pablo Chaves Ávila, and Dimitris Trakas. D1.1–Market and Regulatory Analysis: Analysis of Current Market and Regulatory Framework in the Involved Areas, 2019.

[LC19] Hélène Le Cadre. On the Efficiency of Local Electricity Markets under Decentralized and Centralized Designs: A Multi-Leader Stackelberg Game Analysis. *Central European Journal of Operations Research*, 27(4):953–984, 2019.

[LCJWA20] Hélène Le Cadre, Paulin Jacquot, Cheng Wan, and Clémence Alasseur. Peer-to-peer Electricity Market Analysis: From Variational to Generalized Nash Equilibrium. *European Journal of Operational Research*, 282(2):753–771, 2020.

[LCMP19] Hélène Le Cadre, Ilyès Mezghani, and Anthony Papavasiliou. A Game-Theoretic Analysis of Transmission-Distribution System Operator Coordination. *European Journal of Operational Research*, 274(1):317–339, 2019.

[LL11] Javad Lavaei and Steven H. Low. Zero Duality Gap in Optimal Power Flow Problem. *IEEE Transactions on Power Systems*, 27(1):92–107, 2011.

[LMS94] Irvin J. Lustig, Roy E. Marsten, and David F. Shanno. Interior Point Methods for Linear Programming: Computational State of the Art. *ORSA Journal on Computing*, 6(1):1–14, 1994.

[LPG+19] Guillaume Leclercq, Marco Pavesi, Thomas Gueuning, Araz Ashouri, Peter Sels, Frederik Geth, Reinhilde D'hulst, and Hélène Le Cadre. Network and Market Models. *SmartNet Deliverable D*, 2, 2019.

[LS17] Alvaro Lorca and Xu Andy Sun. The Adaptive Robust Multi-Period Alternating Current Optimal Power Flow Problem. *IEEE Transactions on Power Systems*, 33(2):1993–2003, 2017.

[LW16] Adrian S. Lewis and Stephen J. Wright. A Proximal Method for Composite Minimization. *Mathematical Programming*, 158(1-2):501–546, 2016.

[LWO14] Ruoyang Li, Qiuwei Wu, and Shmuel S. Oren. Distribution Locational Marginal Pricing for Optimal Electric Vehicle Charging Management. *IEEE Transactions on Power Systems*, 29(1):203–211, 2014.

[MAEH99] James A. Momoh, Rambabu Adapa, and Mohamed E. El-Hawary. A Review of Selected Optimal Power Flow Literature to 1993. I. Nonlinear and Quadratic Programming Approaches. *IEEE Transactions on Power Systems*, 14(1):96–104, 1999.

[MEHA99] James A. Momoh, Mohamed E. El-Hawary, and Ramababu Adapa. A Review of Selected Optimal Power Flow Literature to 1993. II. Newton, Linear Programming and Interior Point Methods. *IEEE Transactions on Power Systems*, 14(1):105–111, 1999.

[MGASM+20] Carlos Madina, Inés Gómez-Arriola, Maider Santos-Mugica, Joseba Jimeno, Kris Kessels, Dimitris Trakas, José Pablo Chaves, and Yvonne Ruwaida. Flexibility Markets to Procure System Services. CoordiNet Project. In *2020 17th International Conference on the European Energy Market (EEM)*, pages 1–6. IEEE, 2020.

[MGL14] Kostas Margellos, Paul Goulart, and John Lygeros. On the Road Between Robust Optimization and the Scenario Approach

for Chance Constrained Optimization Problems. *IEEE Transactions on Automatic Control*, 59(8):2258–2263, 2014.

[MH+19] Daniel K. Molzahn, Ian A. Hiskens, et al. *A Survey of Relaxations and Approximations of the Power Flow Equations*. Now Publishers, 2019.

[MHLD13] 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.

[MP19] Ilyes Mezghani and Anthony Papavasiliou. A Mixed Integer Second Order Cone Program for Transmission-Distribution System Co-Optimization. In *2019 IEEE Milan PowerTech*, pages 1–6. IEEE, 2019.

[MPLC18] Ilyès Mezghani, Anthony Papavasiliou, and Hélène Le Cadre. A Generalized Nash Equilibrium Analysis of Electric Power Transmission-Distribution Coordination. In *Proceedings of the Ninth International Conference on Future Energy Systems*, pages 526–531, 2018.

[MR18] Daniel K. Molzahn and Line A. Roald. Towards an AC Optimal Power Flow Algorithm with Robust Feasibility Guarantees. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–7. IEEE, 2018.

[MRH+19] Tillmann Muhlpfordt, Line Roald, Veit Hagenmeyer, Timm Faulwasser, and Sidhant Misra. Chance-Constrained AC Optimal Power Flow: A Polynomial Chaos Approach. *IEEE Transactions on Power Systems*, 34(6):4806–4816, 2019.

[MRS+17] Gianluigi Migliavacca, Marco Rossi, Daan Six, Mario Džamarija, Seppo Horsmanheimo, Carlos Madina, Ivana Kockar, and Juan Miguel Morales. SmartNet: H2020 Project Analysing TSO–DSO Interaction to Enable Ancillary Services Provision from Distribution Networks. *CIRED-Open Access Proceedings Journal*, 2017(1):1998–2002, 2017.

[MSC+17] Salman Mashayekh, Michael Stadler, Gonçalo Cardoso, Miguel Heleno, Sreenath Chalil Madathil, Harsha Nagarajan, Russell Bent, Marc Mueller-Stoffels, Xiaonan Lu, and Jianhui Wang. Security-Constrained Design of Isolated Multi-Energy Microgrids. *IEEE Transactions on Power Systems*, 33(3):2452–2462, 2017.

[MVV15] Mehdi Madani and Mathieu Van Vyve. Computationally Efficient MIP Formulation and Algorithms for European Day-Ahead Electricity Market Auctions. *European Journal of Operational Research*, 242(2):580–593, 2015.

[Nas51] John Nash. Non-Cooperative Games. *Annals of mathematics*, pages 286–295, 1951.

[Nes07] Yurii Nesterov. Modified Gauss–Newton Scheme with Worst Case Guarantees for Global Performance. *Optimisation Methods and Software*, 22(3):469–483, 2007.

[NI+55] Hukukane Nikaidô, Kazuo Isoda, et al. Note on Non-Cooperative Convex Games. *Pacific Journal of Mathematics*, 5(Suppl. 1):807–815, 1955.

[NMRB18] Yeesian Ng, Sidhant Misra, Line A. Roald, and Scott Backhaus. Statistical Learning for DC Optimal Power Flow. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–7. IEEE, 2018.

[NP06] Yurii Nesterov and Boris T. Polyak. Cubic Regularization of Newton Method and Its Global Performance. *Mathematical Programming*, 108(1):177–205, 2006.

[NS20] N-SIDE. Study – System Balancing Solutions with Detailed Grid Data - Phase 2 – Core Analysis, 2020. `https://www.statnett.no/contentassets/3b981e22e5d64179bb22ea9e5b46f515/2020-study---system-balancing-solutions-with-detailed-grid-data.pdf`, [Online].

[NTF11] Koichi Nabetani, Paul Tseng, and Masao Fukushima. Parametrized Variational Inequality Approaches to Generalized Nash Equilibrium Problems with Shared Constraints. *Computational Optimization and Applications*, 48(3):423–452, 2011.

[NW06] Jorge Nocedal and Stephen Wright. *Numerical Optimization*. Springer Science & Business Media, 2006.

[OSH+05] Richard P. O'Neill, Paul M. Sotkiewicz, Benjamin F. Hobbs, Michael H. Rothkopf, and William R. Stewart Jr. Efficient Market-Clearing Prices in Markets with Nonconvexities. *European Journal of Operational Research*, 164(1):269–285, 2005.

[Ost16] Alexander M. Ostrowski. *Solution of Equations and Systems of Equations: Pure and Applied Mathematics: A Series of Monographs and Textbooks, Vol. 9*, volume 9. Elsevier, 2016.

[otNsa18] TSOs of the Nordic synchronous area. Amended Nordic Synchronous Area Proposal for the FRR Dimensioning Rules in Accordance with Article 157(1) of the Commission Regulation (EU) 2017/1485 of 2 August 2017 Establishing a Guideline on Electricity Transmission System Operation, 2018. URL: `https://consultations.entsoe.eu/markets/nordic-tsos-proposals-for-frr-and-frce-and-frequen/supporting_documents/3.%20FRR%20dimensioning%20rules.pdf`.

[PAK16] Ignacio Pérez-Arriaga and Christopher Knittel. Utility of the Future. Technical report, MIT Energy Initiative, December 2016.

[PBDS20] Anthony Papavasiliou, Mette Bjørndal, Gerard Doorman, and Nicolas Stevens. Hierarchical Balancing in Zonal Markets. In *2020 17th International Conference on the European Energy Market (EEM)*, pages 1–6. IEEE, 2020.

[PM18] Anthony Papavasiliou and Ilyes Mezghani. Coordination Schemes for the Integration of Transmission and Distribution System Operations. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–7. IEEE, 2018.

[PSFW08] Jong-Shi Pang, Gesualdo Scutari, Francisco Facchinei, and Chaoxiong Wang. Distributed Power Allocation with Rate Constraints in Gaussian Parallel Interference Channels. *IEEE Transactions on Information Theory*, 54(8):3471–3489, 2008.

[RA17] Line Roald and Göran Andersson. Chance-constrained AC Optimal Power Flow: Reformulations and Efficient Algorithms. *IEEE Transactions on Power Systems*, 33(3):2906–2918, 2017.

[RMKA16] Line Roald, Sidhant Misra, Thilo Krause, and Göran Andersson. Corrective Control to Handle Forecast Uncertainty: A Chance Constrained Optimal Power Flow. *IEEE Transactions on Power Systems*, 32(2):1626–1637, 2016.

[ROKA13] Line Roald, Frauke Oldewurtel, Thilo Krause, and Göran Andersson. Analytical Reformulation of Security Constrained Optimal Power Flow with Probabilistic Constraints. In *2013 IEEE Grenoble Conference*, pages 1–6. IEEE, 2013.

[Ros65] J. Ben Rosen. Existence and Uniqueness of Equilibrium Points for Concave N-person Games. *Econometrica: Journal of the Econometric Society*, pages 520–534, 1965.

[SA74] Brian Stott and Ongun Alsac. Fast Decoupled Load Flow. *IEEE Transactions on Power Apparatus and Systems*, (3):859–869, 1974.

[SBG+20] Bartolomeo Stellato, Goran Banjac, Paul Goulart, Alberto Bemporad, and Stephen Boyd. OSQP: An Operator Splitting Solver for Quadratic Programs. *Mathematical Programming Computation*, pages 1–36, 2020.

[SJA09] Brian Stott, Jorge Jardim, and Ongun Alsaç. DC Power Flow Revisited. *IEEE Transactions on Power Systems*, 24(3):1290–1300, 2009.

[SM79] Brian Stott and J. L. Marinho. Linear Programming for Power-System Network Security Applications. *IEEE Transactions on Power Apparatus and Systems*, (3):837–848, 1979.

[SM20] Tim Schittekatte and Leonardo Meeus. Flexibility Markets: Q&A with Project Pioneers. *Utilities policy*, 63:101017, 2020.

[SOAS10] Yves Smeers, Giorgia Oggioni, Elisabetta Allevi, and Siegfried Schaible. Generalized Nash Equilibrium and Market Coupling in the European Power System. 2010.

[Sto74] Brian Stott. Review of Load-Flow Calculation Methods. *Proceedings of the IEEE*, 62(7):916–929, 1974.

[SV06] Paul M. Sotkiewicz and Jesus M. Vignolo. Nodal Pricing for Distribution Networks: Efficient Pricing for Efficiency Enhancing DG. *IEEE Transactions on Power Systems*, 21(2):1013–1014, 2006.

[SZZL15] Dane A. Schiro, Tongxin Zheng, Feng Zhao, and Eugene Litvinov. Convex Hull Pricing in Electricity Markets: Formulation, Analysis, and Implementation Challenges. *IEEE Transactions on Power Systems*, 31(5):4068–4075, 2015.

[TDD11] Quoc Tran-Dinh and Moritz Diehl. Proximal Methods for Minimizing the Sum of a Convex Function and a Composite Function. Tech. report, KU Leuven, OPTEC and ESAT/SCD, Belgium, May 2011.

[TDGMD11] Quoc Tran-Dinh, Suat Gumussoy, Wim Michiels, and Moritz Diehl. Combining Convex–Concave Decompositions and Linearization Approaches for Solving BMIs, with Application to Static Output Feedback. *IEEE Transactions on Automatic Control*, 57(6):1377–1390, 2011.

[VHM+17] Andreas Venzke, Lejla Halilbasic, Uros Markovic, Gabriela Hug, and Spyros Chatzivasileiadis. Convex Relaxations of Chance Constrained AC Optimal Power Flow. *IEEE Transactions on Power Systems*, 33(3):2829–2841, 2017.

[VMLA13] Maria Vrakopoulou, Kostas Margellos, John Lygeros, and Göran Andersson. A Probabilistic Framework for Reserve Scheduling and Security Assessment of Systems With High Wind Power Penetration. *IEEE Transactions on Power Systems*, 28(4):3885–3896, 2013.

[VVL14] Remco A. Verzijlbergh, Laurens J. De Vries, and Zofia Lukszo. Renewable Energy Sources and Responsive Demand. Do We Need Congestion Management in the Distribution Grid? *IEEE Transactions on Power Systems*, 29(5):2119–2128, September 2014.

[Wäc02] Andreas Wächter. *An Interior Point Algorithm for Large-Scale Nonlinear Optimization with Applications in Process Engineering*. PhD thesis, PhD thesis, Carnegie Mellon University, 2002.

[WB06] Andreas Wächter and Lorenz T. Biegler. On the Implementation of an Interior-Point Filter Line-Search Algorithm for Large-Scale Nonlinear Programming. *Mathematical Programming*, 106(1):25–57, 2006.

[WJ+08] Martin J. Wainwright, Michael I. Jordan, et al. Graphical Models, Exponential Families, and Variational Inference. *Foundations and Trends® in Machine Learning*, 1(1–2):1–305, 2008.

[YAO08] Jian Yao, Ilan Adler, and Shmuel S. Oren. Modeling and Computing Two-settlement Oligopolistic Equilibrium in a Congested Electricity Network. *Operations Research*, 56(1):34–47, 2008.

[YH17] Zhao Yuan and Mohammad Reza Hesamzadeh. Hierarchical Coordination of TSO-DSO Economic Dispatch Considering Large-Scale Integration of Distributed Energy Resources. *Applied Energy*, 195:600–615, 2017.

[ZMST10] 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, 2010.

[ZT11] Baosen Zhang and David Tse. Geometry of Feasible Injection Region of Power Networks. In *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1508–1515. IEEE, 2011.