MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

"Optimization of Remedial Actions in the European Electrical Transmission System using Interior-point Methods"

verfasst von / submitted by

Konstantin Heinrich Sonntag, B.Sc.

angestrebter akademischer Grad / in partial fulfillment of the requirements for the degree of

Master of Science (MSc)

Wien, 2021 / Vienna, 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 821
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Mathematik
Betreut von / Supervisor:	Univ.-Prof. Dr. Radu Ioan Boţ

Abstract / Zusammenfassung

Abstract

In the operation of electrical transmission systems different so-called remedial actions get used by transmission system operators to guarantee operational security of the power grid while securing the power supply at any time. In this thesis, we develop a model to optimize the day-ahead operation of remedial actions for a full day in the European power grid. The remedial actions we consider are the redispatch of ordered generator dispatch, the operation of phase-shifting transformers, and the power flow control on high-voltage DC lines. The optimization model relies on the framework of mathematical linear programming, which gets introduced in this thesis. The general goal of the optimization is to simultaneously minimize power overloads on line elements and the costs that are incurred due to redispatch. We focus on the region Italy North containing the countries Austria, France, Italy, Slovenia, and Switzerland. For this area, we compare three different configurations of our optimization model to analyze the benefits of non-costly remedial like phase-shifting transformers and high-voltage DC lines over the pure use of redispatch. We further introduce a method for the regularization of the activation of phase-shifting transformers, which leads to a more sustainable operation.

Deutsche Zusammenfassung

Übertragungsnetzbetreiber nutzen unterschiedliche sogenannte Remedial Actions, um die Betriebssicherheit des Übertragungsnetzes sicherzustellen und zu jedem Zeitpunkt die Energieversorgung zu gewährleisten. In dieser Abschlussarbeit entwickeln wir ein Modell, um den Einsatz von Remedial Actions für einen gesamten Tag im europäischen Übertragungsnetz zu optimieren. Die Remedial Actions, die wir betrachten, sind der Redispatch des geplanten Dispatches von Kraftwerken, der Einsatz von Phase-Shifting Transformatoren und die direkte Kontrolle des Lastflusses auf Hochspannungs-Gleichstrom Leitungen. Das Optimierungsmodell basiert auf der Mathematik linearer Optimierung, die innerhalb dieser Arbeit präsentiert wird. Das Ziel der Optimierung ist die simultane Minimierung von Leitungsüberlastungen und Kosten, die durch den Einsatz von Redispatch entstehen. Der Fokus innerhalb dieser Arbeit liegt auf der Region Italy North, die aus den Ländern Österreich, Frankreich, Italien, Slowenien und der Schweiz besteht. Für dieses Gebiet vergleichen wir drei verschiedene Modellkonfigurationen unseres Optimierungsmodells, um die Vorteile zu analysieren, die durch den Einsatz von non-costly Remedial Actions wie Phase-Shifting Transformatoren und Hochspannungs-Gleichstrom Leitungen gegenüber der alleinigen Verwendung von Redispatch entstehen. Darüber hinaus führen wir eine Methode zur Regulierung des Einsatzes der Phase-Shifting Transformatoren ein, die zu einer nachhaltigeren Verwendung dieser führt.

Acknowledgement

I would like to thank the people who have supported me in the preparation for this thesis.

First of all, I want to express my gratitude towards my supervisor Prof. Dr. Radu Ioan Boţ who brought this topic to my attention and motivated me to participate in this project. I further want to thank him for encouraging my interest in mathematical optimization throughout my studies in Vienna.

My sincere thanks go to my colleagues at the AIT who made the work on this thesis a great experience. My research would have been impossible without the guidance of my supervisors Ksenia and Gerhard who introduced me to the concepts of transmission systems and remedial action optimization. I also would like to acknowledge my fellow master's student Stefan for the many discussions and his valuable advice. I pay special regards to Tara for her help with all organizational matters.

I am profoundly grateful to my family who helped me in pursuing my path in mathematics. I would like to recognize the assistance of my parents Katharina and Joachim, their constant encouragement and their support of my decisions. I thank my brothers Alexander and Maximilian and my sister Anastasia for the great times we share together.

Special thanks go to my friends in Vienna, back in Germany and all-around Europe for always being there even when I am far off. Last but certainly not least I thank Louisa for her unwavering support and her patience that cannot be underestimated.

Danke Euch allen!

Contents

1 Introduction

1.1 Company Cooperation

This master's thesis is written in cooperation with the *Austrian Institute of Technology GmbH* and the *Austrian Power Grid GmbH*. In this subsection, we give a brief introduction to both companies and their interest in this thesis project.

1.1.1 Austrian Institute of Technology GmbH

This master's thesis is written in cooperation with the *Austrian Institute of Technology GmbH (AIT)*. The AIT is Austria's largest research and technology organization and has has about 1.400 employees mostly at the main locations Vienna Giefinggasse, Seibersdorf, Wiener Neustadt, Ranshofen and Graz. The AIT engages in research in the areas of energy, low-emission transport, mobility systems, bioresources and health, digital safety and security, automation and control and technology experience. The AIT works at the interface of science and industry and enables innovation through its scientific expertise, market experience, customer relationships and research infrastructure [1]. With this thesis project, the AIT aims to achieve a better understanding of the common methodology for regional operational security coordination that gets currently developed by European transmission system operators and the remedial action optimization that is an important component of this methodology. The models that are used in these methodologies are still under active development and the model we build helps to get a better insight into the challenges and opportunities of remedial action optimization.

1.1.2 Austrian Power Grid GmbH

The *Austrian Power Grid GmbH (APG)* is Austria's transmission system operator. The main task of the APG is to secure power supply within the Austrian transmission system by balancing power generation and power consumption. Over 600 employees work at the APG to maintain power balance at every time. In order to realize this, the APG manages many tasks like grid operation, power exchanges with other countries on the electricity market and regulation. The APG maintains a grid with a total length of 6,965 km of different voltage levels. 2,583 km of the grid have a voltage level of 380 kV, 3,206 km have a voltage level of 220 kV and 1,176 km have a voltage level of 110 kV. The grid has 64 substations in total [2]. The main interest of the APG in this thesis relies upon the prevailing development of a common methodology for regional operational security coordination between European transmission system operators. An important component of this methodology is the remedial action optimization that we investigate within this thesis.

1.2 Motivation

In this section, we motivate the problem of remedial action optimization that we investigate in this master's thesis. We present the motivation in two parts. First, we present the so-called *day-ahead markets in electricity*. During the operation of the day-ahead markets certain procedures that mainly focus on market aspects lead to problems in the operation of the electricity grid. In the second part, we describe the *remedial action optimization* that aims to resolve these problems.

1.2.1 The Day-ahead Market in Electricity

In the *European Economic Area*, the internal market in electricity has been gradually established since 1999. The implementation of the internal market in electricity aims at achieving the following objectives [3], [4]:

- securing supply and sustainability of energy
- creating competitive affordable prices
- increasing efficiency
- providing incentives to invest in proper generation (e.g. renewable energy sources)
- decarbonization
- increasing standard of service
- creating new business opportunities
- supporting isolated members in the union

The cooperation between transmission system operators, nominated electricity market operators, and regulatory authorities is necessary to achieve the previous stated goals and to further promote completion and efficient functioning of the internal market in electricity. Single day-ahead and intraday coupling processes require to find compromises between potentially competing power exchanges and are necessary to establish common market coupling functions. Ensuring close cooperation needs a robust, reliable, and non-discriminatory framework for single day-ahead and intraday coupling. This is not possible without a set of unified rules for capacity calculation, congestion management and trading of electricity [5], [6].

Important components of the internal market in electricity are the so-called day-ahead markets. On the pan-European day-ahead markets electricity gets traded in a daily blind auction for every hour of the following day. During this daily auction bids and offers get matched by an algorithm that maximizes social welfare [7]. Because this algorithm does not take the full physical properties of the transmission systems into account, it is possible that the ordered generator dispatch for the next day leads to overloads on line elements. Therefore, additional coordination measures are necessary to guarantee operational security. An important coordination measure is the remedial action optimization that we introduce in the following subsection.

1.2.2 Remedial Action Optimization

In this subsection we describe the *remedial action optimization*. The content of this subsection relies on [5], [8], and [9]. A *remedial action* is any measure a *transmission system operator (TSO)* applies to maintain *operational security*. Operational security is attained when the transmission system is within its acceptable operating boundaries. These boundaries are, for example, thermal limits, voltage limits, frequency, and dynamic stability limits. Further, these boundaries take possible line outages into account that may occur during grid operation.

The remedial action optimization is defined in the *regional operational security coordination methodology (ROSC methodology)* [9]. We do not describe the full regulatory details of the ROSC methodology but focus on the subject matter of remedial action optimization on its own. The ROSC methodology handles both the *day-ahead* and the *intraday* framework. Day-ahead operation means the organization of coordination measures for the next day based on forecasts while intraday operation describes the operation in real-time in reaction to certain events. In this master's thesis, we focus on the day-ahead perspective and remedial action optimization describes the procedure of identifying suitable remedial actions in order to maintain operational security for the next day based on forecasts.

As described in the previous Subsection 1.2.1 the algorithm that matches bids and offers during the market clearing of the day-ahead market in electricity can produce ordered generator dispatch that leads to overloads on transmission lines. Remedial action optimization in the day-ahead framework aims at resolving these overloads to maintain operational security. Using the ordered generator dispatch that gets computed during the market clearing, power flows in the grid can be predicted and possible overloads can be anticipated. Then, remedial action for the next day can be planned in order to resolve congestions.

Remedial actions can be seperated into two classes. On the one hand side *costly remedial actions* and on the other hand side *non-costly remedial actions*. Costly remedial actions are for example the redispatch of ordered generation at a power plant, which involves financial costs. Non-costly remedial actions on the other hand are measures like the use of different pre-installed transformers. In this thesis we want to investigate to what extend the use of non-costly remedial action benefits the remedial action optimization with respect to costs and operational security. We describe the costly and non-costly remedial actions that we consider in this thesis in Chapter 3 and 4. In Chapter 4 we further define how we compare costs and operational security for different configurations of the remedial action optimization.

1.3 Outline

Chapter 2 of this master's thesis contains the mathematical foundation of the concepts that we use to solve the problem of remedial action optimization. In Section 2.1 we introduce the mathematical concept of linear programming which yields a rich framework to formulate many problems in applied areas like engineering, physics, biology, and chemistry. After that, in Subsection 2.2, we introduce the concepts of optimality and duality which lay out the cornerstone for many concepts and algorithms to solve optimization problems. The notions of optimality and duality get applied in Subsection 2.3 of this master's thesis in order to introduce the interior-point method. The interior-point method is our method of choice to solve the problem of remedial action optimization. In Subsection 2.4 we discuss the so-called path-following interior-point method which forms a special instance of the general interior-point method that we have seen before.

Chapter 3 is devoted to transmission systems and the background in electrical engineering. In Section 3.1 we describe the concept of high-voltage AC transmission and formulate the main equations of power flow in a transmission system. Additionally, Sections 3.2 and 3.3 contain the concepts of high-voltage DC transmission and phase-shifting transformers. These form two options of non-costly remedial actions in the remedial action optimization that we investigate in this master's thesis. Section 3.4 serves as a description of the most important distribution factors that are used in power system analysis. Particularly, we use them to identify critical line outages.

In Chapter 4 we present the model we develop for remedial action optimization and discuss our results. In Section 4.1 we give an overview of the most important programming languages and software packages we use. Thereafter, in Section 4.2 the data sources that we use for our model get presented and further we describe in which parts of the model we use the data. Section 4.3 contains the specification of the optimization model with all important equations and constraints. Further, we describe the different model configurations that we compare in this thesis to compare the different optimions in the remedial action optimization. Finally, in Section 4.4, we discuss the results of the optimization and analyze the effects of the different model configurations.

We close this master's thesis with the discussion in Chapter 5. In this chapter we summarize our methodology and findings and draw conclusions from the results to further interpret the effect of the different types of remedial actions on the remedial action optimization.

2 Linear Programming

Linear programming is one of the most powerful tools in numerical mathematics. In practical applications, many problems are formulated as linear programs because for these problems efficient algorithms to derive solutions exist. Therefore, linear optimization problems appear among all other classes of optimization problems most frequently in practice. In Section 4 of this thesis, linear optimization algorithms get applied to solve a load flow problem arising in electrical transmission system operation.

There are two prominent classes of algorithms to solve linear optimization problems. On the one hand, the so-called *simplex method*, originally introduced by Dantzig in 1947 [10] and which was the predominant method of choice for a long time. On the other hand, there is the class of *interior-point methods* which are a reasonable alternative to the simplex method nowadays. In this thesis, we focus on the second approach and give a few remarks on the simplex method within the following sections.

This chapter contains a brief overview of the field of linear programming and beyond that, an introduction to interior-point methods. In Section 2.1 we introduce the linear optimization problems that are investigated in this chapter. Thereafter, in Section 2.2 we describe the so-called *duality theory* for linear programs where we mainly derive important optimality conditions for optimization problems. The optimality conditions help us to understand under which conditions an optimization problem has a solution and they get further employed to derive explicit solution methods. After this, we describe a general interior-point method and a special instance of this method, the so-called *path-following interior-point method* in Sections 2.3 and 2.4.

The content of this Section rests on Chapters 1 to 5 of the book "Theorie und Numerik restringierter Optimierungsaufgaben" by Geiger and Kanzow [11]. We do not give proofs for the theorems in the following sections but refer the interested reader to the book by Geiger and Kanzow and the specified chapters.

2.1 The Optimization Problem

Let $X \subset \mathbb{R}^n$ be a non-empty set and $f : \mathbb{R}^n \to \mathbb{R}$ a given function defined on $\mathbb{R}^n$. Then the general *minimization problem* reads as

$$\text{minimize } f(x) \ \text{ subject to } \ x \in X \tag{2.1}$$

or equivalently written as

$$\begin{array}{ll} \min & f(x) \\ s.t. & x \in X \end{array} .$$

We call the function f the *objective function* of (2.1) and the set X the *feasible set* of (2.1). Analogously, we can formulate the general *maximization problem*

$$\text{maximize } f(x) \ \text{ subject to } \ x \in X. \tag{2.2}$$

Both of these problems are referred to as *optimization problems.*

Definition 2.1.1.
Consider the minimization problem (2.1).

i) A vector $x \in \mathbb{R}^n$ is called feasible *for (2.1), if $x \in X$ holds.*

ii) A feasible vector x^* is called a global minimum *of (2.1), if*

$$f(x^*) \leq f(x) \text{ for all } x \in X$$

holds.

iii) A feasible vector x^* is called a local minimum *of (2.1), if there exists $\varepsilon > 0$ with*

$$f(x^*) \leq f(x) \text{ for all } x \in X \cap \mathcal{U}_\varepsilon(x^*)$$

holds, where $\mathcal{U}_\varepsilon(x^*) := \{x \in \mathbb{R}^n \mid \|x - x^*\| < \varepsilon\}$ is the ε-ball around x^*.

iv) A feasible vector x^* is called a strict global minimum *of (2.1), if*

$$f(x^*) < f(x) \text{ for all } x \in X \text{ with } x \neq x^*$$

holds.

v) A feasible vector x^* is called a strict local minimum *of (2.1), if there exists $\varepsilon > 0$ such that*

$$f(x^*) < f(x) \text{ for all } x \in X \cap \mathcal{U}_\varepsilon(x^*) \text{ with } x \neq x^*$$

holds.

Analogously to Definition 2.1.1 we could define the notion of (strict) global and local maxima. But since x^* is a global (or local) maximum of (2.2) if and only if x^* is a global (or local) minimum of

$$\begin{array}{rl} \min & -f(x) \\ s.t. & x \in X \end{array}$$

we do not need this definition. This is the reason why we can focus on minimization problems without loss of generality when we study optimization problems.
Using Definition 2.1.1, we can precisely describe what we mean by solving the minimization problem (2.1), namely as finding a global (or local) minimum x^* of (2.1). Usually, the task of finding a global minimum is hard and, therefore, in many areas of optimization we are interested in finding local minima. In this thesis, we focus on linear optimization problems. When studying these, we are in the fortunate situation that local minima coincide with global minima, since linear optimization problems are a special instance of convex optimization problems.
Oftentimes, the feasible set $X \subset \mathbb{R}^n$ is defined using functions $g : \mathbb{R}^n \to \mathbb{R}^m$ and $h : \mathbb{R}^n \to \mathbb{R}^p$. To use these functions, when defining the feasible set we need to extend the inequality to a componentwise notion on $\mathbb{R}^n$ as we do in the following definition.

Definition 2.1.2.
For $x, y \in \mathbb{R}^n$ we define the relations

$$x \leq y :\Longleftrightarrow x_i \leq y_i \textit{ for all } i = 1, \ldots, n,$$

and

$$x < y :\Longleftrightarrow x_i < y_i \textit{ for all } i = 1, \ldots, n.$$

In this context, we denote the zero element of any finite dimensional Euclidean space $\mathbb{R}^n$ with 0 and write

$$x \leq 0 :\Longleftrightarrow x_i \leq 0 \textit{ for all } i = 1, \ldots, n,$$

and

$$x < 0 :\Longleftrightarrow x_i < 0 \textit{ for all } i = 1, \ldots, n.$$

Using the notation from the previous definition, the constraint $g(x) \leq 0$ reads as

$$g(x) \leq 0 \iff g_i(x) \leq 0 \text{ für alle } i = 1, \ldots, m.$$

We then can define the feasible set

$$X = \{x \in \mathbb{R}^n \mid g(x) \leq 0,\ h(x) = 0\}. \tag{2.3}$$

When using a feasible set given by functions $g : \mathbb{R}^n \to \mathbb{R}^m$ and $h : \mathbb{R}^n \to \mathbb{R}^p$ as in (2.3), we write the minimization problem (2.1) as

$$\begin{array}{rr} \min & f(x) \\ s.t. & x \in \mathbb{R}^n, \\ & g(x) \leq 0, \\ & h(x) = 0. \end{array} \tag{2.4}$$

The importance of problems of the form (2.1) and (2.4) stems from the fact that they are *mathematical models* for many real-world problems in engineering, physics, chemistry, medicine, economics, and ecology. The components $x \in \mathbb{R}^n$ are parameters of the model that are not fixed in advance and are free to be chosen. For these parameters optimal solutions are desired. In Chapter 1 of [11], various examples for applied optimization problems are stated. In Section 4 of this thesis, we will further see that this premise holds when we apply optimization algorithms to solve an optimal load flow problem that arises in electrical transmission system operation.

For optimization problems of the form (2.1) and (2.4) we need to deepen the investigations of the problem's structure if we want to derive methods that systematically find solutions. For arbitrary functions f, g, and h, we cannot hope to find a global minimum with an algorithm. We need further information on properties of the functions like continuity, differentiability, convexity, or linearity. In this thesis, we are interested in optimization problems of the form (2.4) with the function f being linear while the functions g and h are affine transformations. This means we can write the functions f, g, and h as

$$\begin{aligned} f(x) &= c^T x && \text{for} \quad c \in \mathbb{R}^n, \\ g(x) &= Cx - d && \text{for} \quad C \in \mathbb{R}^{p \times n} \text{ and } d \in \mathbb{R}^p, \\ h(x) &= Ax - b && \text{for} \quad A \in \mathbb{R}^{m \times n} \text{ and } b \in \mathbb{R}^m. \end{aligned}$$

Then the linear optimization problem (2.4) reads as

$$\begin{array}{rr} \min & c^T x \\ s.t. & x \in \mathbb{R}^n, \\ & Cx \leq d, \\ & Ax = b. \end{array} \tag{2.5}$$

A linear optimization problem of the form (2.5) can always be transformed into an equivalent linear problem in *standard form* that we introduce in the following definition.

Definition 2.1.3.
For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$ define the linear optimization problem in standard form as

$$\begin{array}{rr} \min & c^T x \\ s.t. & x \in \mathbb{R}^n, \\ & Ax = b, \\ & x \geq 0. \end{array} \tag{2.6}$$

In this definition, the number of inequality and equality constraints does not align with n and m as in (2.4). Nevertheless we stick to the formulation of the linear problem in standard form as it is most convenient in literature.
By equivalent linear optimization problems we mean that every feasible point of one optimization problem corresponds to a feasible point of the other optimization problem and that an optimal solution of one optimization problem corresponds in the same way to an optimal solution of the other optimization problem. To give a better understanding of this equivalence to the reader, we will transform the problem (2.5) into the form of problem (2.6). Start by looking at

$$\begin{array}{rr} \min & c^T x \\ s.t. & x \in \mathbb{R}^n, \\ & Cx \leq d, \\ & Ax = b. \end{array}$$

We can transform the inequality $Cx \leq d$ in an equality using a non-negative slack variable $s \in \mathbb{R}^p$ to get the problem

$$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & x \in \mathbb{R}^n,\ s \in \mathbb{R}^p, \\ & Cx + s = d, \\ & Ax = b, \\ & s \geq 0. \end{aligned}$$

To substitute x by non-negative variable we define $x^+, x^- \in \mathbb{R}^n$ with $x^+, x^- \geq 0$ and set $x = x^+ - x^-$. Doing so, the optimization problem reads as

$$\begin{aligned} \min \quad & c^T (x^+ - x^-) \\ s.t. \quad & x^+, x^- \in \mathbb{R}^n,\ s \in \mathbb{R}^p, \\ & C(x^+ - x^-) + s = d, \\ & A(x^+ - x^-) = b, \\ & s, x^+, x^- \geq 0. \end{aligned}$$

We can finally transform this into the standard form by rewriting the equations in matrix form as

$$\begin{aligned} \min \quad & \begin{pmatrix} c \\ -c \\ 0 \end{pmatrix}^T \begin{pmatrix} x^+ \\ x^- \\ s \end{pmatrix} \\ s.t. \quad & \begin{pmatrix} x^+ \\ x^- \\ s \end{pmatrix} \in \mathbb{R}^{2n+p}, \\ & \begin{pmatrix} C & -C & \mathrm{Id} \\ A & -A & 0 \end{pmatrix} \begin{pmatrix} x^+ \\ x^- \\ s \end{pmatrix} = \begin{pmatrix} d \\ b \end{pmatrix} \\ & \begin{pmatrix} x^+ \\ x^- \\ s \end{pmatrix} \geq 0, \end{aligned}$$

where $\mathrm{Id} \in \mathbb{R}^{p \times p}$ is the identity matrix and $0 \in \mathbb{R}^{m \times p}$ is the zero matrix. This problem is written in standard form. All optimization variables are non-negative and we only have linear equality constraints. This transformation justifies that one can study optimization problems that are written in standard form without loss of generality.

The feasible set for the optimization problem (2.6) is

$$X := \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\}.$$

The simplex method relies on the geometrical properties of this set. In fact, X is a convex polyhedron and one can show that under certain conditions (that are in practice often met) an optimal solution to (2.6) is a vertex of the polyhedron X. Therefore, the strategy of the simplex method is to find a path along the edges of the polyhedron

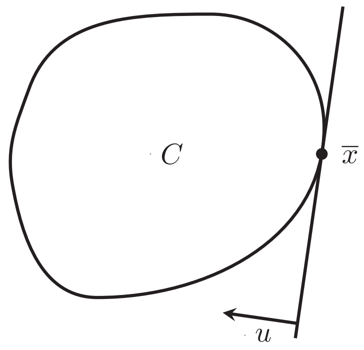


Figure 2.1: Separation Lemma

X to the optimal vertex. The interior-point methods, as their name suggests, find a path to the optimal solution through the interior of the set X as we see in the later subsections. Consequently, we do not further investigate all the geometrical properties of X as they are not necessary to introduce the interior-point methods. On the other hand, to derive the interior-point methods we need some background in the optimality theory of nonlinear optimization. We will give background on this in the following subsection.

2.2 Duality and Optimality

2.2.1 Separation Theorems

In this subsection, we present the separation theorem and the strict separation theorem. Speaking visually, these theorems show that under certain conditions two convex sets can be (strictly) separated by a hyperplane. These statements, as fundamental as they seem, are yet powerful and form the foundation of duality theory. Further, the proof of the *Farkas Lemma*, which gets presented in the next subsection, relies on the strict separation theorem.

Lemma 2.2.1.
Let $C \subset \mathbb{R}^n$ non-empty and convex and let $\overline{x} \in \mathbb{R}^n$ a vector not belonging to the interior of C. Then there exists $u \in \mathbb{R}^n \setminus \{0\}$ that satisfies

$$u^T \overline{x} \leq u^T x, \quad \forall x \in C.$$

This lemma states that for a non-empty and convex set C and a point $\overline{x}$ that does not belong to the set C, a hyperplane that separates the point and the set exists. The proof relies on convex analysis and the existence of the projection on a closed convex set. Figure 2.1 visualizes Lemma 2.2.1.
It is obvious that the inequality of Lemma 2.2.1 also holds for all elements in the closure $\mathrm{cl}\,(C)$ of C. If we choose $\gamma := u^T \overline{x}$, then the inequality states that all elements $x \in C$

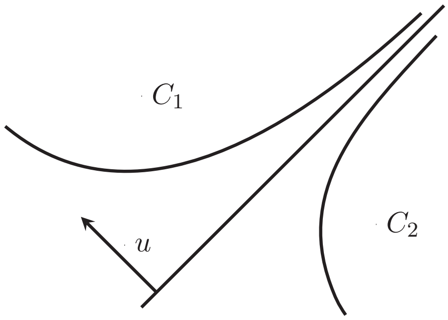


Figure 2.2: Separation Theorem

belong to one of the half-spaces that get separated by the hyperplane

$$H := \left\{x \in \mathbb{R}^n \mid u^T x - \gamma = 0\right\},$$

namely $H_{\leq} := \left\{x \in \mathbb{R}^n \mid u^T x - \gamma \leq 0\right\}$ and $H_{\geq} := \left\{x \in \mathbb{R}^n \mid u^T x - \gamma \geq 0\right\}$. The vector u is a normal vector to the hyperplane H. Using Lemma 2.2.1, we are in the position to prove the *separation theorem.*

Theorem 2.2.2. *(Separation Theorem)*
Let $C_1, C_2 \subset \mathbb{R}^n$ be two non-empty and disjoint convex sets. Then, there exists a vector $u \in \mathbb{R}^n \setminus \{0\}$ that satisfies

$$u^T x_1 \leq u^T x_2, \quad \forall x_1 \in C_1, x_2 \in C_2.$$

The proof of this theorem relies on the fact that the set

$$C := C_2 - C_1 := \{x \in \mathbb{R}^n \mid x = x_2 - x_1 \text{ for some } x_1 \in C_1, x_2 \in C_2\}$$

is non-empty and convex. Since C_1 and C_2 are disjoint, it follows that zero is not an element of C. Using Lemma 2.2.1 on the set C and the zero vector yields the desired result. Theorem 2.2.2 is visualized in Figure 2.2.
We want to improve this statement further. To this end, we need an additional lemma. Before we formulate this next lemma, we want to remark that the sum of two non-empty, convex and closed sets in general is not closed itself. To derive the closedness of the sum one needs additional assumptions on the sets that get presented in the following lemma.

Lemma 2.2.3.
Let $C_1, C_2 \subset \mathbb{R}^n$ two non-empty and convex sets with C_1 closed and C_2 compact. Then, the sum $C_1 + C_2 := \{x \in \mathbb{R}^n \mid x = x_1 + x_2 \text{ for some } x_1 \in C_1, x_2 \in C_2\}$ is also a non-empty, convex and closed set.

In the proof, we look at a convergent sequence $\left\{x_1^k + x_2^k\right\}$. Without the additional compactness assumption there is no way to deduce the convergence of the sequences

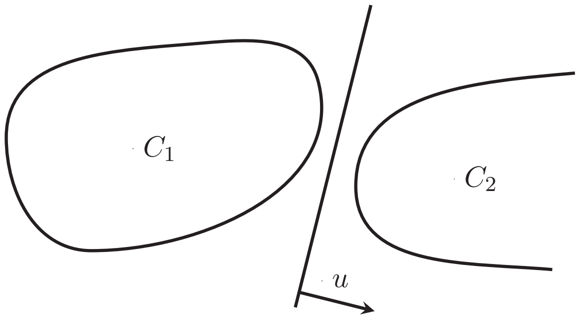


Figure 2.3: Strict separation Theorem

$\{x_1^k\}$ and $\{x_2^k\}$. For unbounded sets C_1 and C_2 the sequences can be unbounded but their convergence is necessary to proof that $\{x_1^k + x_2^k\}$ converges inside $C_1 + C_2$. Using this lemma, we can proceed with the *strict separation theorem*.

Theorem 2.2.4. *(Strict separation theorem)*
Let $C_1, C_2 \subset \mathbb{R}^n$ be two non-empty, disjoint and convex sets, where C_1 is closed and C_2 is compact. Then, there exist a non-trivial vector $u \in \mathbb{R}^n$ and a scalar $\gamma \in \mathbb{R}$ that satisfy

$$u^T x_1 < \gamma < u^T x_2, \quad \forall x_1 \in C_1,\ x_2 \in C_2.$$

The proof of this can be seen as a refinement of the proof of Theorem 2.2.2 and relies solely on convex analysis. Once again, the compactness assumption is crucial. Without this it is possible that the two convex sets get asymptotically closer to the hyperplane that separates them and, therefore, to one another. Figure 2.2 hints at this problem. Here, the sets C_1 and C_2 get asymptotically closer. In Figure 2.3 the set C_1 is compact meaning this problem does not arise.
The strict separation theorem can be used to proof Farkas Lemma that we present in the following subsection.

2.2.2 Farkas Lemma

In this subsection, we introduce the famous Farkas Lemma. Farkas Lemma can be proven efficiently using the Strict Separation Theorem 2.2.4. It is of major importance in optimality theory and in the duality theory of linear programming. Before we move to Farkas Lemma, we have to define cones and investigate their properties.

Definition 2.2.5.
A subset $K \subset \mathbb{R}^n$ is a cone, if

$$\lambda x \in K$$

holds, for all $x \in X$ and for all $\lambda > 0$.

We call a cone $K \subset \mathbb{R}^n$ a convex cone, closed cone, etc., if K is a closed, convex, etc. subset of $\mathbb{R}^n$.
Let $a_1, \dots, a_n \in \mathbb{R}^n$ be given vectors, then the set

$$\text{cone}\{a_1, \dots, a_n\} \coloneqq \{x_1 a_1 + \cdots + a_n x_n \mid x_i \geq 0,\ \forall i = 1, \dots, n\}$$

is a convex cone in $\mathbb{R}^n$. It is called the *cone generated by* $a_1, \dots, a_n$. The following result gives insight into the properties of such a cone. While the statement in the Lemma looks trivial the proof is quite technical.

Lemma 2.2.6.
Let $A \in \mathbb{R}^{m\times n}$ be given. Then, the set

$$X \coloneqq \left\{y \in \mathbb{R}^n \mid y = A^T x,\ x \geq 0, \right\}$$

is a non-empty, closed and convex cone.

That the set X is a non-empty, and convex cone follows directly from its definition. Showing that X is closed is the hard part of the proof that gets quite technical. Using Lemma 2.2.6, we can now prove the celebrated Farkas Lemma.

Lemma 2.2.7. *(Farkas Lemma)*
Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Then, the following statements are equivalent:

a) The system $A^T x = b$, $x \geq 0$ has a solution.

b) The inequality $b^T d \geq 0$ holds for all $d \in \mathbb{R}^m$ with $Ad \geq 0$.

While the proof of the direction $a) \Rightarrow b)$ follows immediately, the proof of the other direction is more subtle. The proof relies on the fact that $X \coloneqq \left\{y \in \mathbb{R}^n \mid y = A^T x,\ x \geq 0\right\}$ is a non-empty, closed, and convex cone due to Lemma 2.2.6 and on the strict separation theorem applied to the set X and the compact, convex set $\{b\}$.
We mainly use Farkas Lemma in Subsection 2.2.4 to prove optimality conditions for the linear optimization problem (2.6). Before doing so, we introduce the *Karush-Kuhn-Tucker conditions*, likewise showing an indirect use of Farkas Lemma.

2.2.3 The Karush-Kuhn-Tucker Conditions

In this subsection, we introduce optimality conditions for nonlinear optimization problems. We avoid unnecessary detail by only stating the very results we need for deriving the interior-point method.
Again, we look at the optimization problem (2.4) that reads as

$$\begin{array}{rr} \min & f(x) \\ s.t. & x \in \mathbb{R}^n, \\ & g(x) \leq 0, \\ & h(x) = 0, \end{array}$$

with continuously differentiable functions f, g and h.

Definition 2.2.8.
The function $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ that is defined by

$$L(x, \lambda, \mu) \coloneqq f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{p} \mu_j h_j(x)$$

is called the Lagrange function of the optimization problem (2.4).

Using the Lagrange function, we can define the so-called *Karush-Kuhn-Tucker conditions* for the optimization problem (2.4).

Definition 2.2.9.
Consider the optimization problem (2.4) with continuously differentiable functions f, g, and h.

i) The conditions

$$\begin{aligned} \nabla_x L(x, \lambda, \mu) &= 0 \\ h(x) &= 0 \\ \lambda \geq 0,\ g(x) \leq 0,\ \lambda^T g(x) &= 0 \end{aligned}$$

are called Karush-Kuhn-Tucker conditions (KKT conditions) of the optimization problem (2.4), where

$$\nabla_x L(x, \lambda, \mu) = \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) + \sum_{j=1}^{p} \mu_j \nabla h_j(x)$$

is the gradient of the Lagrange function L with respect to the x variable.

ii) Every vector $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$, that satisfies the KKT conditions, is called a Karush-Kuhn-Tucker point (KKT point) of the optimization problem (2.4). The vectors λ^* and μ^* are called Lagrange multipliers.

We define the following optimization problem with linear constraints,

$$\begin{array}{ll} \min & f(x) \\ s.t. & x \in \mathbb{R}^n \\ & a_i^T x \leq \alpha_i,\ i = 1, \ldots, m, \\ & b_j^T x = \beta_j,\ j = 1, \ldots, p, \end{array} \tag{2.7}$$

which is an instance of the optimization problem (2.4), with

$$g_i(x) \coloneqq a_i^T x - \alpha_i,\ i = 1, \ldots, m \ \text{ and } \ h_j(x) \coloneqq b_j^T x - \beta_j,\ j = 1, \ldots, p.$$

Furthermore, let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function, $a_i, b_j \in \mathbb{R}^n$ given vectors, and $\alpha_i, \beta_j \in \mathbb{R}$ given scalars for $i = 1, \ldots, m$, $j = 1, \ldots, p$.
The next result states that the KKT conditions for a problem of this type are always necessary optimality conditions without the use of additional constraint qualifications. To put it in other words, the fact that the constraints are described by linear functions is a constraint qualification on its own.

Theorem 2.2.10. *(KKT conditions for linear constraints)*
Let x^* be a local minimum of the optimization problem (2.7). Then, there exist Lagrange multipliers $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^p$ such that the vector (x^*, λ^*, μ^*) satisfies the KKT conditions

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* a_i + \sum_{j=1}^{p} \mu_j^* b_j = 0, \\ b_j^T x^* = \beta_j \; j = 1, \dots, p, \\ a_i^T x^* \leq \alpha_i \; i = 1, \dots, m, \\ \lambda_i^* (a_i^T x^* - \alpha_i) = 0 \; i = 1, \dots, m, \\ \lambda_i^* \geq 0 \; i = 1, \dots, m \end{aligned}$$

for problem (2.7).

The proof of this theorem relies on the Abadie constraint qualification that is formulated using the tangential cone of the feasible set. We do not further introduce these concepts as we merely want to give background to introduce the interior-point method.

In Section 2.3, we define auxiliary problems that are not linear but convex. To handle these problems we need certain optimality conditions for convex optimization problems that rely on the KKT conditions. Therefore, in the remainder of this subsection we state the most important optimality conditions for convex optimization problems using the KKT conditions.
The convex optimization problem reads as

$$\begin{array}{rrl} \min & f(x) & \\ s.t. & x \in \mathbb{R}^n & \\ & g_i(x) \leq 0, & i = 1, \dots, m, \\ & b_j^T x = \beta_j, & j = 1, \dots, p, \end{array} \tag{2.8}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$ are continuously differentiable and convex functions and $b_j \in \mathbb{R}^n$ are given vectors and $\beta_j \in \mathbb{R}$ are given scalar values for $j = 1, \dots, p$. So, we are looking at an optimization problem with a convex objective function, convex inequality constraints and linear equality constraints. The feasible set of such a problem is convex since it is the intersection of convex sets and affine linear subspaces. Therefore, we call the minimization problem (2.8) a *convex optimization problem.* We discuss the important properties of convex optimization problems in the remainder of this subsection.

Lemma 2.2.11.
Consider the optimization problem

$$\begin{array}{ll} \min & f(x) \\ s.t. & x \in X \end{array}, \tag{2.9}$$

where $X \subset \mathbb{R}^n$ is a convex set and $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function. Then, every local minimum of the convex optimization problem (2.8) is also a global minimum.

The proof of this lemma follows immediately by the definition of convexity. For convex optimization problems a useful constrained qualification exists, that we introduce with the following definition.

Definition 2.2.12. *(Slater constrained qualification)*
The convex optimization problem (2.8) satisfies the Slater constrained qualification (Slater CQ), if there exists a vector $\hat{x} \in \mathbb{R}^n$ that fulfills

$$g_i(\hat{x}) < 0, \;\; i = 1, \ldots, m \;\;\; \textit{and} \;\;\; b_j^T \hat{x} = \beta_j, \;\; j = 1, \ldots, p,$$

namely $\hat{x}$ is strictly feasible with respect to the inequality constraints and feasible with respect to the equality constraints.

The next result states that for a local (= global) minimum of a convex optimization problem (2.8) satisfying the Slater constrained qualification there exist Lagrange multipliers.

Theorem 2.2.13. *(KKT conditions under Slater CQ)*
Let $x^* \in \mathbb{R}^n$ be a local (= global) minimum of the convex optimization problem (2.8). Let the Slater constraint qualification be fulfilled. Then, there exist Lagrange multipliers $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^p$ such that, the vector $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ satisfies the KKT conditions

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^{m} \lambda^* \nabla g_i(x^*) + \sum_{j=1}^{p} \mu_j^* b_j = 0, \\ b_j^T x^* = \beta_j, j = 1, \ldots, p, \\ g_i(x^*) \leq 0, j = 1, \ldots, m, \\ \lambda_i^* g_i(x^*) = 0, j = 1, \ldots, m, \\ \lambda_i^* \geq 0, j = 1, \ldots, m, \end{aligned} \tag{2.10}$$

of (2.8).

It also holds that for the convex optimization problem (2.8) the validity of the KKT conditions is already a sufficient optimality condition, as we will state in the following result. The following theorem does also hold without the Slater constrained qualification which is of major importance to proof Theorem 2.2.13.

Theorem 2.2.14.
Let $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ be a KKT point of the convex optimization problem (2.8). Then, x^* is a local (= global) minimum of the convex optimization problem (2.8).

A direct consequence of Theorem 2.2.10 and Theorem 2.2.14 is the following result on the connection between the KKT points and the solutions of a convex optimization problem with linear constraints.

Corollary 2.2.15.
Consider the linear optimization problem 2.8, with all constraints given by linear functions. Then $x^* \in \mathbb{R}^n$ is a local (= global) minimum of (2.8), if there exist Lagrange multipliers $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^p$ such that the vector (x^*, λ^*, μ^*) is a KKT point of (2.8).

Corollary 2.2.15 is of major importance as we use it to investigate the optimality conditions of the linear optimization problems in the following subsection. It is used in Subsection 2.3 to derive interior-point methods to solve linear optimization problems.

2.2.4 Duality

In this subsection, we return to linear optimization problem in standard form (2.6) that we denote from here onwards with (P). For $A \in \mathbb{R}^{m\times n}$, $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ the problem (P) reads as

$$\begin{array}{rr} \min & c^T x \\ s.t. & x \in \mathbb{R}^n, \\ & Ax = b, \\ & x \geq 0. \end{array} \tag{P}$$

This problem is closely related to the problem

$$\begin{array}{rr} \max & b^T \lambda \\ s.t. & \lambda \in \mathbb{R}^m, \\ & A^T\lambda \leq c, \end{array}$$

as we see throughout this subsection. Introducing a non-negative slack variable, we get the following reformulation of the problem:

$$\begin{array}{rr} \max & b^T \lambda \\ s.t. & \lambda \in \mathbb{R}^m, s \in \mathbb{R}^n, \\ & A^T\lambda + s = c, \\ & s \geq 0. \end{array} \tag{D}$$

Note that (D) is again a linear optimization problem. Both, the minimization problem (P) and the maximization problem (D), are defined using the data (A, b, c). The problem (P) is called the *primal linear program*, while the problem (D) is called the *dual linear program*.

The goal of this subsection is to discuss the relation between the primal problem (P) and the dual problem (D). We also want to show that the introduction of (D) is sensible. A first important result is stated in the following theorem.

Theorem 2.2.16. *(Optimality conditions)*
The following statements are equivalent

(a) The primal problem (P) has a solution x^*

(b) The dual problem (D) has a solution (λ^*, s^*)

(c) The so-called system of optimality conditions

$$\begin{aligned} A^T\lambda + s &= c, \\ Ax &= b, \\ x_i s_i &= 0, \ \forall i = 1, \ldots, n, \\ x, s &\geq 0, \end{aligned} \tag{2.11}$$

has a solution (x^*, λ^*, s^*).

The proof of Theorem 2.2.16 follows from Corollary 2.2.15 and the fact that the KKT conditions of the optimization problem (P) are equivalent to *(c)*. The same holds for the dual problem (D).

Theorem 2.2.17. *(Weak duality)*
Let $x \in \mathbb{R}^n$ *be a feasible point of the primal problem (P) and let* $(\lambda, s) \in \mathbb{R}^m \times \mathbb{R}^n$ *be a feasible point of the dual problem (D). Then, it holds that*

$$b^T\lambda \leq c^T x.$$

Proof:
The proof of Theorem 2.2.17 follows immediately from the feasibility of x and (λ, s). In fact, the proof can be written in one line as

$$b^T\lambda = (Ax)^T\lambda = x^T(A^T\lambda) = x^T(c - s) \leq c^T x,$$

where the last inequality follows from the fact that $x^T s \geq 0$ which is a consequence of $x \geq 0$ and $s \geq 0$. □

From Theorem 2.2.17 it follows that every feasible point of the dual program (D) yields a lower bound for the optimal function value of the primal problem (P). Vice versa every feasible point of the primal problem yields an upper bound for the optimal value of the dual objective function. We introduce the notations

$$\inf(P) := \inf\left\{c^T x \mid Ax = b, \ x \geq 0\right\}$$

and

$$\sup(D) := \sup\left\{b^T\lambda \mid A^T\lambda + s = c, \ s \geq 0\right\}$$

for the optimal function values of the primal and dual problem respectively. From the weak duality theorem it directly follows that

$$\sup(D) \leq \inf(P).$$

Here we set

$$\inf(P) := +\infty \quad \text{and} \quad \sup(D) := -\infty$$

respectively if the primal or the dual problem do not have any feasible points.
A direct consequence of Theorem 2.2.17 is the following existence result for linear optimization problems.

Corollary 2.2.18.
Let $x^* \in \mathbb{R}^n$ be a feasible point of the primal problem (P) and $(\lambda^*, s^*) \in \mathbb{R}^m \times \mathbb{R}^n$ a feasible point of the dual problem (D), that fulfill

$$c^T x^* = b^T \lambda^*.$$

Then, x^* is a solution to the primal problem (P) and (λ^*, s^*) a solution to the dual problem (D).

The proof directly follows from Theorem 2.2.17. Another consequence of weak duality is that

$$c^T x - b^T \lambda \geq 0$$

for all primal feasible points x and all dual feasible points (λ, s). Due to Corollary 2.2.18, we know that for the primal and the dual problem there exist solutions if $c^T x - b^T \lambda = 0$ holds. In this case, we have $\inf(P) = \sup(D)$, while in general we only have $\inf(P) \geq \sup(D)$. If $\inf(P) > \sup(D)$ we say there is a *duality gap*. Also, the value of $c^T x - b^T \lambda$ is sometimes called the duality gap.
With the next theorem we show that the reversal of Corollary 2.2.18 holds as well, namely that from the existence of a solution to the primal problem (P) and the dual problem (D) we can deduce that there is no duality gap. This statement is called the *strong duality theorem.*

Theorem 2.2.19. *(Strong duality)*
If the primal problem (P) has a solution x or the dual problem (D) has a solution (λ, s), then $\inf(P) = \sup(D)$, and the duality gap vanishes.

Proof:
The proof of this statement relies on Theorem 2.2.16 and the fact that for a vector $(x, \lambda, s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ satisfying the conditions (2.11) the equality

$$0 = s^T x = (c - A^T \lambda)^T x = c^T x - \lambda^T (Ax) = c^T x - b^T \lambda \tag{2.12}$$

follows. □

Theorem 2.2.19 states that the primal problem (P) has a solution if and only if the dual problem (D) has a solution. But all together it is still not clear which conditions lead to the existence of a solution to one of these problems. The following theorem gives a sufficient condition. The proof of the theorem relies heavily on Farkas Lemma 2.2.7.

Theorem 2.2.20. *(Existence theorem)*
The following statements hold:

(a) If $\inf(P) \in \mathbb{R}$*, then the primal problem (P) has a solution.*

(b) If $\sup(D) \in \mathbb{R}$*, then the dual problem (P) has a solution.*

If the feasible set for the primal problem (P) and the dual problem (D) is not empty, then the weak duality according to Theorem 2.2.17 gives

$$-\infty < \sup(D) \leq \inf(P) < +\infty.$$

Then, both $\sup(D)$ and $\inf(P)$ are finite and Theorem 2.2.20 leads to the following corollary.

Corollary 2.2.21.
If both, the primal and the dual linear program are feasible, then both problems have an optimal solution.

2.3 Interior-point Methods

In this section, we introduce the basic ideas of behind interior-point methods. Subsection 2.3.1 is devoted to the central path that is important for the general understanding of the interior-point methods. The general interior-point method that we introduce in Subsection 2.3.3 relies on Newton's method which we describe in Subsection 2.3.2.

2.3.1 The Central Path

Throughout this subsection we focus on the primal linear optimization problem

$$\begin{array}{rr} \min & c^T x \\ s.t. & x \in \mathbb{R}^n, \\ & Ax = b, \\ & x \geq 0, \end{array} \tag{P}$$

and its corresponding dual problem

$$\begin{array}{rr} \max & b^T \lambda \\ s.t. & \lambda \in \mathbb{R}^m, \\ & A^T \lambda \leq c, \end{array} \tag{D}$$

for $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. Due to Theorem 2.2.16, the primal linear program and the dual linear program are related to the optimality conditions

$$\begin{aligned} A^T\lambda + s &= c, \\ Ax &= b, \\ x_i s_i &= 0, \ \forall i = 1, \ldots, n, \\ x, s &\geq 0. \end{aligned} \tag{2.13}$$

For one of these problems there exists a solution if and only if there exists a solution for one of the other problems. This observation justifies the approach to solve the primal and the dual problem using the optimality conditions (2.13). The methods presented in this subsection rely on this idea. For this reason, they are called *primal-dual methods*, since they make use of the primal variables x and of the dual variables (λ, s). *Primal methods* and *dual methods* that make use solely of the primal or dual variables exist as well. In general, the purely primal and purely dual methods are numerically inferior to the primal-dual methods and not further discussed here.
In this subsection, we are interested in the following perturbation of the optimality conditions (2.13):

$$\begin{aligned} A^T\lambda + s &= c, \\ Ax &= b, \\ x_i s_i &= \tau, \ \forall i = 1, \ldots, n, \\ x, s &> 0, \end{aligned} \tag{2.14}$$

where $\tau > 0$ is a given parameter. We call this perturbed optimality conditions the *central path conditions.* If $(x_\tau, \lambda_\tau, s_\tau)$ is a (hopefully unique) solution to (2.14), we call the mapping

$$\tau \mapsto (x_\tau, \lambda_\tau, s_\tau)$$

the *central path.* The central path plays a major role for almost all interior-point methods because the fundamental idea behind most of these is to numerically follow the central path more or less exactly.
At this moment, it is not clear if the central path conditions have a solution $(x_\tau, \lambda_\tau, s_\tau)$ for all or even one $\tau > 0$.
In this subsection, we show that under certain assumptions on the conditions (2.14) the existence and uniqueness of solutions follows. To this end, we introduce for the primal linear program (P) the corresponding logarithmic barrier-problem

$$\begin{array}{ll} \min & c^T x - \tau \sum_{i=1}^n \log(x_i) \\ s.t. & x \in \mathbb{R}^n, \\ & Ax = b, \\ & x > 0. \end{array} \tag{PB}$$

The barrier-problem corresponding to the dual problem (D) then is:

$$
\begin{array}{ll}
\max & b^T\lambda + \tau \sum_{i=1}^{n} \log(s_i) \\
s.t. & \lambda \in \mathbb{R}^m, \\
& A^T\lambda + s = c, \\
& s > 0.
\end{array}
\tag{DB}
$$

Our next result is similar to Theorem 2.2.16 and states that the central path condition (2.14) has a unique solution when one of the primal and the dual barrier-problems (PB) and (DB) has a solution. To proof this statement, we use Corollary 2.2.15. Therefore, consider the optimization problem

$$
\begin{array}{ll}
\min & f(x) \\
s.t. & x \in \mathbb{R}^n, \\
& Ax = b, \\
& x \in X,
\end{array}
\tag{2.15}
$$

with a continuously differentiable and convex objective function $f : \mathbb{R}^n \to \mathbb{R}$, a matrix $A \in \mathbb{R}^{m \times n}$, a vector $b \in \mathbb{R}^m$ and $X \subset \mathbb{R}^n$ an open and convex set. The problem (2.15) is a convex optimization problem with linear equality constraints and with an additional open set as a constraint. Since the constraints that are introduced by X are always inactive, the statement of Corollary 2.2.15 still holds for the problem (2.15). This means a vector $x^* \in X$ is a solution to (2.15) if there exist Lagrange multipliers $\lambda^* \in \mathbb{R}^m$ such that $(x^*, \lambda^*) \in X \times \mathbb{R}^m$ satisfy the KKT conditions

$$
\begin{aligned}
\nabla f(x) + A^T\lambda &= 0, \\
Ax &= b.
\end{aligned}
$$

We use this version of Corollary 2.2.15 in the proof of the following result.

Theorem 2.3.1.
Let $\tau > 0$. Then, the following statements hold:

a) The primal barrier-problem has a solution x_τ.

b) The dual barrier-problem has a solution (λ_τ, s_τ).

c) The central path conditions have a solution $(x_\tau, \lambda_\tau, s_\tau)$.

To prove the existence of a solution to the central path conditions (2.14), it suffices by Theorem 2.3.1 to show the existence of a solution to the primal barrier-problem (PB). Before we exploit this fact, we want to recap that the central path conditions not always have a solution. Therefore, at this point, we need an additional condition. To this end, we define the *primal-dual feasible set*

$$
\mathcal{F} := \left\{ (x, \lambda, s) \mid Ax = b,\ A^T\lambda + s = c,\ x, s \geq 0 \right\}
$$

and the *primal-dual strict feasible set*

$$\mathcal{F}^o := \big\{(x,\lambda,s) \mid Ax = b,\ A^T\lambda + s = c,\ x, s > 0\big\}$$

that we call from this point on *feasible set* and *strict feasible set*. The set $\mathcal{F}$ (and $\mathcal{F}^o$) contains exactly the points $(x,\lambda,s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$, which (strictly) satisfy the constraints of the primal linear program (P) and the constraints of the dual linear program (D).
If $(x_\tau, \lambda_\tau, s_\tau)$ is a solution to the central path conditions (2.14) for $\tau > 0$, then this vector is an element of the strict feasible set $\mathcal{F}^o$. If this set is empty, there cannot exist any solutions to the central path conditions (2.14). Likewise, due to Theorem 2.3.1, the primal barrier-problem (PB) then does not possess any solutions. Therefore, $\mathcal{F}^o \neq \emptyset$ is a necessary condition for the existence of an optimal solution to (PB). The next result states that this condition is also sufficient.

Theorem 2.3.2.
Let the strict feasible set $\mathcal{F}^o$ be non-empty. Then, the primal barrier-problem for every $\tau > 0$ has a solution x_τ.

Using this theorem, we are in the position to state the main result of this subsection.

Theorem 2.3.3. *(Existence of the central path)*
Let the strict feasible set $\mathcal{F}^o$ be non-empty. Then, for every $\tau > 0$ the central path condition has a solution $(x_\tau, \lambda_\tau, s_\tau)$ and the x and s components of this vector are uniquely determined. If the matrix A has full rank, then the λ components are uniquely determined as well.

The interior-point method that we discuss in the following subsections is an application of Newton's method to the system of equations in the central path conditions. Therefore Newton's method is of major importance for the deduction of the interior-point method and we dedicate the following subsection to a brief summary of Newton's method.

2.3.2 A Recap of Newton's Method

The content of this subsection relies on Chapter 5 of [11]. Let $F : \mathbb{R}^p \to \mathbb{R}^p$ be a continuously differentiable function. We are interested in finding a solution $w^* \in \mathbb{R}^p$ to the nonlinear system of equations

$$F(w) = 0. \tag{2.16}$$

Probably, the most infamous method to solve such a system of equations is Newton's method. We briefly describe it in this subsection.
Assume we have an approximation w^k of the solution of (2.16) that we denote with w^*. To compute an even better approximation w^{k+1}, we linearize the function F at the point w^k. We denote this linearization with

$$F_k(w) := F(w^k) + F'(w^k)(w - w^k).$$

Since the nonlinear problem $F(w) = 0$ is hard to solve, we approximate it in the k-th iteration with the linear system of equations

$$F_k(w) = 0.$$

The solution of this system gives the next iterate w^{k+1}. From the definition of F_k we get the explicit description

$$w^{k+1} = w^k - F'(w^k)^{-1}F(w^k), \tag{2.17}$$

if the inverse $F'(w^k)^{-1}$ does exist. Obviously, we are not interested in computing the inverse matrix $F'(w^k)^{-1}$ explicitly but computing the correction $d^k \in \mathbb{R}^p$ by solving the so called *Newton equation*

$$F'(w^k)d = -F(w^k)$$

and then set

$$w^{k+1} = w^k + d^k. \tag{2.18}$$

This produces the same vector w^{k+1} as in (2.17). This method gets iterated (as long as $F'(w^k)$ is regular) and results in a sequence $\{w^k\}_{k\geq 0}$ that hopefully converges to w^*. This procedure to iteratively solve a nonlinear system of equation in the form of (2.16) is called Newton's method. We formalize this in the following algorithm.

Algorithm 2.3.4 Newton's method

1: Choose $w^0 \in \mathbb{R}^p$, and set $k = 0$
2: **if** $F(w^k) = 0$ **then**
3: STOP
4: **end if**
5: Compute $d^k \in \mathbb{R}^p$ a solution to the Newton equation

$$F'(w^k)d = -F(w^k)$$

6: Set $w^{k+1} := w^k + d^k$, $k \leftarrow k + 1$ and go to step 2

At this point we do neither know if Algorithm 2.3.4 converges at all nor if it is well-defined. The next theorem states a theoretical result that describes the power of Algorithm 2.3.4.

Theorem 2.3.5.
Let $F : \mathbb{R}^p \to \mathbb{R}^p$ be continuously differentiable, $w^* \in \mathbb{R}^p$ be a zero of F and the Jacobian $F'(w^*)$ be regular. Then, there exists $\varepsilon > 0$ such that for all $w^0 \in \mathcal{U}_\varepsilon(w^*)$ it holds that:

a) Newton's method as defined in algorithm 2.3.4 is well-defined and generates a sequence $\left\{w^k\right\}_{k\geq 0}$ that converges to w^*.

b) The rate of convergence is superlinear.

c) If F' is locally Lipschitz-continuous, then the rate of convergence is quadratic.

Superlinear convergence in part *b)* means that there exists a sequence $\{\varepsilon^k\}_{k\geq 0} \subset \mathbb{R}_{\geq 0}$ that converges to zero such that for all $k \geq 0$ it holds that $\left\|w^{k+1} - w^*\right\| \leq \varepsilon_k \left\|w^k - w^*\right\|$. Quadratic convergence in *c)* is defined by the existence of a positive constant C such that $\left\|w^{k+1} - w^*\right\| \leq C \left\|w^k - w^*\right\|^2$ holds for all $k \geq 0$.

Theorem 2.3.5 states that for a suitable starting vector w^0 Newton's method exhibits fast convergence to a solution of the nonlinear system of equations (2.16). While the conditions on the function F in practice are often met, the condition on the starting vector is the major theoretical drawback of Newton's method. To guarantee convergence, we must choose a starting vector in a neighborhood of the solution w^* that is potentially very small. For an arbitrary starting vector the behavior of the sequence generated by Newton's method can be chaotic and there is in general no chance to tell what happens. Although this condition on the starting vector looks like a big drawback, Newton's method and its modifications are used by practitioners with big success. A possible way to globalize Newton's method is to define suitable step sizes $t_k > 0$ and change the update rule (2.18) to

$$w^{k+1} := w^k + t_k d^k.$$

If we have an almost singular Jacobian $F'(w^k)$ it can happen that the step direction d^k has a huge norm. Choosing a small t^k in such situations helps the algorithm to not overshoot in critical situations. We use this globalization in the following subsection.

2.3.3 Basics of Interior-point Methods

In this subsection, we apply a globalized Newton's method to the equations in the central path conditions (2.14). The additional inequalities that arise in the central path condition get neglected for a moment. Also, the interior-point methods in general do not fulfill the condtions that are stated in Theorem 2.3.5 to guarantee convergence. But this should not be of any concern for now.
In order to apply Newton's method to the central path conditions (2.14), we define for $\tau > 0$ the function F_τ as

$$F_\tau(w) := F_\tau(x, \lambda, s) := \begin{pmatrix} A^T\lambda + s - c \\ Ax - b \\ XSe - \tau e \end{pmatrix}, \tag{2.19}$$

with the notations

$$\begin{aligned} X &:= \operatorname{diag}(x_1, \ldots, x_n), \\ S &:= \operatorname{diag}(s_1, \ldots, s_n), \\ e &:= (1, \ldots, 1)^T. \end{aligned}$$

Using F_τ the central path conditions (2.14) are equivalent to

$$F_\tau(x, \lambda, s) = 0,\ x > 0,\ s > 0. \tag{2.20}$$

We are now able to apply Algorithm 2.3.4 to the system $F_\tau(w) = 0$. The additional strict inequalities from (2.20) get ignored for now. To apply Newton's method to this problem, we investigate if the Jacobian

$$F'(x, \lambda, s) = \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix}$$

is regular under certain assumptions. We state this within the following result.

Theorem 2.3.6.
Let $w = (x, \lambda, s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ be a vector with $x > 0$ and $s > 0$. The matrix $A \in \mathbb{R}^{m\times n}$ may have full rank. Then, the Jacobian $F'_\tau(w)$ is regular for all $\tau > 0$.

Notice that the conditions $x > 0, s > 0$ in Theorem 2.3.6 are not crucial, because all interior-point methods in general only produce positive iterates x^k and s^k. In fact, the regularity of the Jacobian is one of the reasons why interior-point methods produce positive iterates.
Also, the condition that the matrix A has full rank is not critical. From a theoretical point of view we can assume that A has full rank. Numerically this is more complicated even though in practice most methods try to remove redundant restriction from A by making use of clever linear algebra.
Theorem 2.3.6 justifies the application of Newton's method to (2.19) the system of equations given by $F_\tau(w) = 0$. For a given iteration vector $w^k = (x^k, \lambda^k, s^k)$, we can set $w^{k+1} = w^k + t_k \Delta w^k$ for a certain step size $t_k > 0$ and a correction $\Delta w^k = (\Delta x^k, \Delta\lambda^k, \Delta s^k)$ that is a solution to the Newton equation

$$F'_{\tau_k}(w^k)\Delta w = -F_{\tau_k}(w^k).$$

By the definition of F_τ in (2.19) this is equivalent to

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \\ \Delta s \end{pmatrix} = \begin{pmatrix} -A^T\lambda^k - s^k + c \\ -Ax^k + b \\ -X^k S^k e + \tau_k e \end{pmatrix}, \tag{2.21}$$

with

$$X^k := \operatorname{diag}(x_1^k, \dots, x_n^k) \ \text{ and } \ S^k := \operatorname{diag}(s_1^k, \dots, s_n^k).$$

We want to see what happens when the k-th iterate $w^k = (x^k, \lambda^k, s^k)$ fulfills the equations from the central path condition, namely $A^T\lambda^k + s^k = c$ and $Ax^k = b$ holds. From the first equation in block form in the Newton equation (2.21) it follows that

$$A^T\Delta\lambda^k + \Delta s^k = 0.$$

Because $w^{k+1} = w^k + t_k \Delta w^k$ we can deduce that

$$\begin{aligned} A^T\lambda^{k+1} + s^{k+1} - c &= A^T(\lambda^k + t_k\Delta\lambda^k) + (s^k + t_k\Delta s^k) - c \\ &= A^T\lambda^k + s^k - c + t_k(A^T\Delta\lambda^k + \Delta s^k) \\ &= 0. \end{aligned}$$

Analogously, from the second equation in block form of (2.21) it follows that

$$Ax^{k+1} = Ax^k + t_k A\Delta x^k = b.$$

Therefore, as $w^k = (x^k, \lambda^k, s^k)$ also the next iterate $w^{k+1} = (x^{k+1}, \lambda^{k+1}, s^{k+1})$ satisfies the linear equations $A^T\lambda + s = c$ and $Ax = b$. Therefore, all iterates satisfy both of these equations if the starting vector $w^0 = (x^0, \lambda^0, s^0)$ satisfies them. From this it follows that in the Newton equation (2.21) we can substitute the first two blocks on the right hand side by zeros.
After these introductory observations, we now define a very general interior-point method, which uses the strict feasible set

$$\mathcal{F}^o := \left\{(x, \lambda, s) \mid Ax = b,\ A^T\lambda + s = c,\ x, s > 0\right\},$$

which is introduced in Subsection 2.3.1.

Algorithm 2.3.7 General interior-point method

1: Choose $w^0 := (x^0, \lambda^0, s^0) \in \mathcal{F}^0, \varepsilon \in (0, 1)$ and set $k := 0$
2: **if** $\mu_k := (x^k)^T s^k / n \leq \varepsilon$ **then**
3:    STOP
4: **end if**
5: Choose $\sigma^k \in [0, 1]$, and find a solution $\Delta w^k := (\Delta x^k, \Delta\lambda^k, \Delta s^k)$ of the following linear system of equations:

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -X^k S^k e + \sigma_k \mu_k e \end{pmatrix} \tag{2.22}$$

6: Set $w^{k+1} := w^k + t_k\Delta w^k, k \leftarrow k + 1$, and go to step 1. Doing so choose $t_k > 0$ a step size, that satisfies $x^{k+1} > 0$ and $s^{k+1} > 0$.

Before we discuss the convergence properties of Algorithm 2.3.7, we want to give some remarks. Because the starting vector (x^0, λ^0, s^0) is an element of the strict feasible set $\mathcal{F}^0$ and the step size t_k is chosen to ensure the positivity of x^k and s^k, it follows from our previous observations that the vector (x^k, λ^k, s^k) belongs to the set $\mathcal{F}^0$ for all $k \geq 0$. In particular, x^k is always primal feasible and (λ^k, s^k) is always dual feasible. From (2.12) it follows that

$$(x^k)^T s^k = c^T x^k - b^T \lambda^k$$

and therefore the term $(x^k)^T s^k$ in the termination condition of Algorithm 2.3.7 is equal to the duality gap that corresponds to the vectors x^k and λ^k. This is the only part in the Newton equation that we have to reduce to zero.
By Theorem 2.3.6 the linear system of equations (2.22) always has a unique solution (if A has full rank) and we can also find $t_k > 0$ such that for $x^k > 0$ and $s^k > 0$ it also holds that $x^{k+1} > 0$ and $s^{k+1} > 0$. Therefore, Algorithm 2.3.7 is well-defined.

Algorithm 2.3.7 is formulated in a relative general manner with two main degrees of freedom. First the so-called *centering parameter* σ_k and second the step size $t_k > 0$. Depending on the choice of this parameters, we get different interior-point methods. In the next subsection, we focus on a special realization of Algorithm 2.3.7.

We also want to give some additional remarks on the centering-parameter σ_k. The term $\sigma_k \mu_k$ that appears on the right-hand side of (2.22) basically plays the role of the parameter τ in the central path conditions. While μ_k is used to describe the weighted duality gap

$$\mu_k := (x^k)^T s^k / n,$$

we are free in our choice of σ_k. The choice $\sigma_k = 0$ corresponds to a Newton step applied to the optimality conditions (2.13) (and these are the conditions we want to solve). However, the choice of $\sigma_k = 0$ usually results in small step sizes since we have to guarantee the positivity of the iterates x^k and s^k. The other extreme choice is $\sigma_k = 1$ which does not get us closer to a solution of (2.13) but closer to the central path and allows bigger step sizes.

In the next step, we show that the (weighted) duality gap μ_k can get reduced in every step depending on our choices of σ_k and t_k. We formulate the next result for an arbitrary step size $t > 0$, where we use the notations

$$\begin{aligned} \big(x^k(t), \lambda^k(t), s^k(t)\big) &:= \big(x^k, \lambda^k, s^k\big) + t\big(\Delta x^k, \Delta\lambda^k, \Delta s^k\big) \\ \mu_k(t) &:= x^k(t)^T s^k(t)/n. \end{aligned}$$

In the next lemma we investigate the relation between the step sizes Δx^k and $\Delta\mu^k$ and further the relation between μ^k and $\mu^k(t)$.

Lemma 2.3.8.
The solution $\big(\Delta x^k, \Delta\lambda^k, \Delta s^k\big)$ of the linear system of equation (2.22) has the following properties

a) $\big(\Delta x^k\big)^T \Delta s^k = 0$.

b) $\mu_k(t) = (1 - t(1-\sigma_k))\mu_k$.

Within the next theorem we state a result that describes the convergence of Algorithm 2.3.7.

Theorem 2.3.9.
Let $\varepsilon \in (0,1)$ be arbitrary and let $\{(x^k, \lambda^k, s^k)\}_{k\geq 0}$ be a sequence generated by Algorithm 2.3.7. The parameters in Algorithm 2.3.7 may fulfill the condition

$$\mu_{k+1} \leq \left(1 - \frac{\delta}{n^{\omega}}\right)\mu_k, \quad k = 1, 2, \ldots \tag{2.23}$$

for some positive constants δ and ω. If the start vector (x^0, λ^0, s^0) fulfills the condition

$$\mu_0 \leq \frac{1}{\varepsilon^{\kappa}} \tag{2.24}$$

for a positive constant κ, then there exists an index $K \in \mathbb{N}$ with

$$K = O(n^{\omega}|\log(\varepsilon)|)$$

and

$$\mu_k \leq \varepsilon$$

for all $k \geq K$.

The main requirement for Theorem 2.3.9 is the condition (2.23). In the following section, we give an example for an interior-point method that satisfies this condition. The condition (2.24) is a restriction on the choice of the starting vector. If we can prove for an interior-point method that it satisfies the conditions in Theorem 2.3.9, we prove that such a method produces iterates with a vanishing duality gap. We apply this theorem in the following section when we investigate a specific instance of the general interior-point method that is defined in Algorithm 2.3.7.

2.4 Path-Following Interior-point Methods

In this section, we describe a specific instance of Algorithm 2.3.7 to solve linear optimization problems. The idea behind *path-following interior-point methods*, as their name implies, is to numerically follow the central path. Since we cannot follow the path exactly, we will construct a neighborhood of the path that we want to pace and that indirectly measures the inexactness of the method.
One differentiates between *feasible* and *infeasible* path-following methods. For feasible methods the starting vector and all following iterations satisfy the linear equalities $Ax = b$ and $A^T\lambda + s = c$ from the central path conditions, while the infeasible methods may produce iterates that do not satisfy these conditions. In this thesis we will only consider one feasible method. The content of this section again follows [11] and in addition [12].
The method we derive in this section will generate iterates in the following neighborhood $\mathcal{N}_{-\infty}(\gamma)$ of the central path:

$$\mathcal{N}_{-\infty}(\gamma) := \{(x, \lambda, s) \in \mathcal{F}^o | x_i s_i \geq \gamma\mu \text{ for all } i = 1, \ldots, n\},$$

with $\gamma \in (0,1)$ a given parameter and

$$\mu := x^T s/n$$

the (weighted) duality gap corresponding to the vector (x, λ, s). For the boundary case $\gamma = 0$ the set $\mathcal{N}_{-\infty}(\gamma)$ coincides with the set of strict feasible points $\mathcal{F}^o$. Next, we formulate a feasible interior-point method and investigate its properties. Doing so, we use the notations

$$\left(x^k(t), \lambda^k(t), s^k(t)\right) := \left(x^k + t\Delta x^k, \lambda^k + t\Delta\lambda^k, s^k + t\Delta s^k\right)$$

and

$$\mu_k(t) := x^k(t)^T s^k(t)/n.$$

that we already have introduced to formulate Lemma 2.3.8. Using them, we define the following algorithm.

Algorithm 2.4.1 A feasible path-following interior-point method

1: Choose $\gamma \in (0,1)$, $0 < \sigma_{\min} < \sigma_{\max} < 1$, $\varepsilon \in (0,1)$, $w^0 := (x^0, \lambda^0, s^0) \in \mathcal{N}_{-\infty}(\gamma)$, and set $k = 0$.
2: **if** $\mu_k := (x^k)^T s^k/n \le \varepsilon$ **then**
3: STOP
4: **end if**
5: Choose $\sigma^k \in [\sigma_{\min}, \sigma_{\max}]$, and find a solution $\Delta w^k := (\Delta x^k, \Delta\lambda^k, \Delta s^k)$ of the following linear system of equations:

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -X^k S^k e + \sigma_k \mu_k e \end{pmatrix} \tag{2.25}$$

6: Let t_k be the biggest step size $t \in [0,1]$ with

$$\left(x^k(t), \lambda^k(t), s^k(t)\right) \in \mathcal{N}_{-\infty}(\gamma).$$

7: Set $w^{k+1} := w^k + t_k \Delta w^k$, $k \leftarrow k+1$, and go to step 1.

Algorithm 2.4.1 is in fact a special instance of Algorithm 2.3.7. For the centering parameter σ_k in step 2 still remains a certain degree of freedom, but the extreme cases $\sigma_k = 0$ and $\sigma_k = 1$ get excluded uniformly using the lower and upper bound $\sigma_{\min}$ and $\sigma_{\max}$. The choice of the step size $t_k > 0$ in step 2 in the contrary is fixed. The convergence properties of the algorithm do not change if we choose another step size. For example, we can replace the rule for t_k in step 2 by a backtracking strategy. Alternatively, we can use the step size $\bar{t}_k$ that we introduce in Lemma 2.4.4 instead of t_k.

The remainder of this section is devoted to the theoretical investigation of Algorithm 2.4.1. We want to show that for a suitable starting vector the method terminates after finitely many iterations due to the criterion that is imposed in step 1. We further want to show that the number of iterations the algorithm needs until it terminates depends polynomially on the number of variables n. Methods with this property are called methods with *polynomial complexity.* We also want to remark that the simplex method does not have this property. There are known counter examples where the number of iterations for the simplex method grows exponentially with the number of variables n most renown by Klee and Minty [13].
Before we state the main results of this section, we need a few helping results that are quite technical. In the following, we introduce three lemmata to make Theorem 2.3.9 of the last section applicable to Algorithm 2.4.1.

Lemma 2.4.2.
Let $u, v \in \mathbb{R}^p$ be two vectors with $u^T v \geq 0$. Then it holds that

$$\|UVe\| \leq 2^{-3/2} \|u+v\|^2$$

with $U = diag(u_1, \ldots, u_n)$, $V = diag(v_1, \ldots, v_n)$ and $e = (1, \ldots, 1) \in \mathbb{R}^p$.

The proof of this lemma relies on fundamental inequalities and on inequalities between the ℓ_1-norm and the ℓ_2-norm.
In the next result we use the notations

$$\Delta X^k := \text{diag}(\Delta x_1^k, \ldots \Delta x_n^k) \quad \text{and} \quad \Delta S^k := \text{diag}(\Delta s_1^k, \ldots \Delta s_n^k),$$

where $(\Delta x^k, \Delta\lambda^k, \Delta s^k)$ is meant to be a solution to linear system of equation (2.25).

Lemma 2.4.3.
Let $\left(x^k, \lambda^k, s^k\right) \in \mathcal{N}_{-\infty}(\gamma)$ be given. Then

$$\left\|\Delta X^k \Delta S^k e\right\| \leq 2^{-3/2}(1+1/\mu)n\mu_k$$

The next lemma gives a lower bound for the step size t_k. This result is a main ingredient to prove the polynomial complexity of Algorithm 2.4.1.

Lemma 2.4.4.
Let $\left(x^k, \lambda^k, s^k\right) \in \mathcal{N}_{-\infty}(\gamma)$ be given. Then

$$\left(x^k(t), \lambda^k(t), s^k(t)\right) \in \mathcal{N}_{-\infty}(\gamma)$$

for all $t \in [0, \bar{t}_k]$ with

$$\bar{t}_k := 2^{3/2}\gamma\frac{\sigma_k}{n}\frac{1-\gamma}{1+\gamma}$$

Using the lemma above we can give an estimate on the reduction of μ_k in every step.

Theorem 2.4.5.
Let $\left\{\left(x^k, \lambda^k, s^k\right)\right\}_{k\geq 0}$ be a sequence given by Algorithm 2.4.1. Then it holds, that

$$\mu_{k+1} \leq \left(1 - \frac{\delta}{n}\right)\mu_k, \quad k = 0, 1, 2, \ldots$$

for a constant $\delta > 0$ that is independent from k.

The proof of this theorem relies on the previous lemmata. With this theorem, we are finally in the position to prove the main convergence theorem for Algorithm 2.4.1. This theorem states that Algorithm 2.4.1 terminates after $\mathcal{O}\left(n\left|\log(\varepsilon)\right|\right)$ iteration due to the condition in step 1. The constant in the $\mathcal{O}$-statement depends on the quality of the starting vector.

Theorem 2.4.6.
Let $\left\{\left(x^k, \lambda^k, s^k\right)\right\}_{k\geq 0}$ be a sequence generated by Algorithm 2.4.1. Assume the starting vector $\left(x^0, \lambda^0, s^0\right)$ satisfies the condition

$$\mu_0 \leq \frac{1}{\varepsilon^\kappa}$$

for a positive constant κ. Then there exists $K \in \mathbb{N}$ with $K = \mathcal{O}\left(n\left|\log(\varepsilon)\right|\right)$ and

$$\mu_k \leq \varepsilon$$

for all $k \geq K$.

The proof of this theorem follows immediately by combining Theorem 2.4.5 and Theorem 2.3.9.

3 Transmission Systems

In this section, we introduce the basic concepts of transmission systems and load flow calculation. The introduction on transmission system stems from [14]. Electrical energy is transported and distributed mostly via alternating current or more general via *high-voltage AC transmission (HVAC)*. In special cases two buses can be connected through *high-voltage DC transmission (HVDC)*. Both methods get introduced in the following subsections. We also want to emphasize that in this thesis we do not describe the background in electrical engineering in too much detail. The focus of this thesis lies on the mathematical description and application of optimization and the modelling of the problem of remedial action optimization. Therefore, we will not describe how the equations for load flow calculations arise from the physical principles and engineering of transmission lines and transformers but recommend the interested reader to the sources [14], [15], [16], [17], [18] and [19]. These sources contain detailed descriptions of the application of electrical engineering and further details on the physical derivation and modelling.

3.1 High-voltage AC Transmission

Transport, transmission and distribution of electrical energy is structured in hierarchically ordered voltage levels, so called *grids*.
A grid is the entirety of all connected operating resources of the same nominal voltage level. One differentiates between

- *Transmission system* 220 kV, 380 kV, for big distances 765 kV
- *Subtransmission system* with voltages of 110 kV
- *Primary and secondary distribution systems* 10 kV, 20 kV and sometimes 30 kV
- *Low-voltage distribution systems* 235 V, 400 V and 690 V

All levels are coupled over transformers. In this thesis, the focus lies on the transmission system. The subtransmission and distribution systems are not discussed because we only model parts of the European transmission system.

3.1.1 Transmission Systems

Transmission systems transmit large amounts of electrical energy in changing direction within and between different regions and between spatial separated centers of energy production and energy consumption. Examples are the transport of energy from location-bound water or coal-powered power stations into densely populated areas or the transport of low-priced nuclear energy over far distances within the interconnected European power grid. There are also short transport routes for power plant feed-in into the power grid. In general, the transmission system is meshed. This means that every bus is connected by at least two branches. If one branch has an outage the other

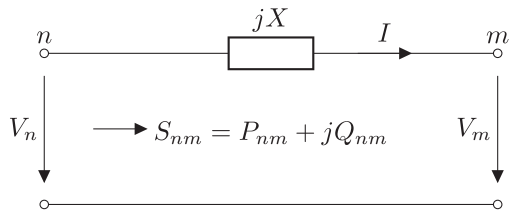


Figure 3.1: Equivalent circuit diagram

branches take over energy supply.

The power flow may change direction because as the day progresses different cheapest producing power plants come into action and therefore the power gets to the centers of consumptions on different paths. The driving force of the energy are power deficiencies in the consumer buses and power surpluses in the generator buses that get balanced by *load flows*. The balancing load flows pass through the grid independent of ownership structures on different parallel paths to the centers of consumption. The distribution of the load flow on the different paths takes place according to Ohm's law for alternating currents depending on the reciprocal of the branch impedances. Without external intervention congestions of branches with low impedances can arise. It is the duty of the *Transmission system operator* to control the power flow by choosing suitable electrical feed, the grid topology and settings of transformers to secure energy supply while preventing congestions and outages.

3.1.2 Power Through a Transmission Line

The content of this subsection relies on [14], [15] and [16]. The load flow on an AC transmission line is determined by the voltage magnitude and phase angle at the start and end bus and by the line reactance. For transmission lines one can neglect the line resistance ($X \gg R$) and setting it to zero does not lead to big errors. In Figure 3.1 one can see the equivalent circuit diagram that models an AC transmission line from bus n to bus m with line reactance X and complex bus voltages V_n and V_m. We are interested in the complex power S and the active power P in particular which is the real part of S. In this chapter we denote with $j \in \mathbb{C}$ the imaginary unit $j^2 = -1$ because this notation is more convenient in electrical engineering in contrast to the mathematical notation $i^2 = -1$. By Kirchhoffs second law $\sum_k V_k = 0$ we can deduce from Figure 3.1 that

$$V_n - V_m - jXI = 0$$

and solve for the complex current to get

$$I = \frac{V_n - V_m}{jX}. \tag{3.1}$$

We write the complex voltages using their magnitude and phase angle as

$$V_n = |V_n|\, e^{j\theta_n} \quad \text{and} \quad V_m = |V_m|\, e^{j\theta_m},$$

and define the line phase angle as the difference of the bus phase angles as

$$\delta := \theta_n - \theta_m.$$

The single-phase complex line power at node n can be obtained by

$$S_{nm} = V_n \overline{I}, \tag{3.2}$$

where $\overline{I}$ is the complex conjugate of I. Substituting (3.1) into (3.2) yields

$$\begin{aligned}
S_{nm} &= V_n \frac{\overline{V_n} - \overline{V_m}}{-jX} \\
&= \frac{V_n \overline{V_n} - |V_n|\, e^{j\theta_n}\, |V_m|\, e^{-j\theta_m}}{-jX} \\
&= \frac{|V_n|^2 - |V_n|\, |V_m|\, e^{j\delta}}{-jX} \\
&= \frac{|V_n|^2 - |V_n|\, |V_m|\, (\cos(\delta) + j\sin(\delta))}{-jX} \\
&= \frac{|V_n|\, |V_m| \sin(\delta)}{X} + j\frac{|V_n|^2 - |V_n|\, |V_m| \cos(\delta)}{X} \\
&= P_{nm} + jQ_{nm}.
\end{aligned}$$

Therefore, the transmitted active and reactive power equal

$$P_{nm} = \frac{|V_n|\, |V_m| \sin(\delta)}{X} \quad \text{and} \quad Q_{nm} = \frac{|V_n|^2 - |V_n|\, |V_m| \cos(\delta)}{X}, \tag{3.3}$$

as functions of the line reactance, the line phase angle and the bus voltage magnitudes. The direction of the power flow depends on the sign of δ and therefore on the bus voltage angles θ_n and θ_m. The relations (3.5) gets used in the following subsection when we introduce a linearization for the active power flow on a line.

3.1.3 DC Load Flow Calculation

This subsection relies on [15], [16], [18] and [19]. The *DC load flow calculation* is a technique to approximate only the active power flows in a power system. If we want to compute the active power flows in a grid using (3.5), we have to solve a highly nonlinear system of equations. In practice this can be done using Newton's method that we introduce in Subsection 2.3.2. When we want to use the load flow equations (3.5) in an optimization model as we do in Chapter 4 the nonlinearities introduce a huge amount of complexity. In practice, it is often not feasible to optimize a model with a full grid in reasonable time using the nonlinear equations. Therefore, one often uses a linearization of the load flow equations. The most common linearization is the so-called DC load flow calculation. The DC load flow calculation uses assumptions on the line reactance, the bus voltages and line phase angles to simplify the load flow equations and give a more direct solution to the load flow problem. The assumptions are the following:

- The voltage profile is assumed to be flat, that means that all voltage magnitudes are close to 1 per unit.
- The X/R ratio is assumed to be large, so that resistances are negligible.
- The bus phase angle differences are assumed to be small. From this it follows that we can approximate $\sin\delta \approx \delta$.

Using these assumptions, we can retrieve the linear relation

$$P_{nm} = \frac{|V_n|\,|V_m|\sin(\delta)}{X} \approx \frac{\theta_n - \theta_m}{X}, \tag{3.4}$$

that we use in particular in Chapter 4 to develop an optimization model for power flow control.

3.2 High-voltage DC Transmission

The information of this section is taken from [14], [17] and [18]. The transmission of huge amounts of energy over far distances of 600 km and more takes place via *high-voltage direct current lines (HVDCs)*. With HVDCs, full active power control is possible, and a HVDC line can be modeled by replacing it by a power injection/withdrawl at the receiving/sending bus. When the losses in the HVDC line are neglected, the injected power is equal to the power withdrawn from the sending bus. Therefore the power flow is not effected by the bus voltage angles at the sending and receiving bus.

3.3 Phase-shifting Transformers

In Subsections 3.1.2 and 3.1.3, we see that the load flow on a line depends on the line reactance, the line phase angle and the bus voltage magnitudes. By controlling these parameters it is possible to change the load flow on a line. A possible measure to change the line phase angles are so-called *phase-shifting transformers* (*PSTs*). A PST is a mechanically device which inserts a transformed variable line voltage in series with the phase voltage using a tap changer. In this thesis, we do not go into detail with respect to the electrotechnical aspects but refer the interested reader to [16], [17] and [18]. In this sources one can find technical descriptions of different PST architectures and also application of PSTs for power flow control. We do mainly focus on the mathematical modeling of the PSTs that we also adopt from these sources.
A PST can be modelled as a reactance X_{PST} in series with an ideal phase shift α.

Applied on equation (3.5) for the transmitted active power this yields

$$P_{nm} = \frac{|V_n|\,|V_m|\sin(\delta+\alpha)}{X + X_{PST}} \tag{3.5}$$

Under the DC load flow assumptions from Subsection 3.1.3 with the additional assumptions that we only allow small phase angles α, we can adopt in the spirit of equation (3.4) the approximation

$$P_{nm} = \frac{|V_n|\,|V_m|\sin(\delta+\alpha)}{X + X_{PST}} \approx \frac{\delta+\alpha}{X + X_{PST}},$$

for the load flow on a line that is equipped with a PST. Again the approximation $\sin(\delta+\alpha) \approx \delta+\alpha$ is only valid for small angles. At an angle of 30°, this simplification leads to an error of about 5% for the active power flow on the line.

3.4 Distribution Factors

In power system analysis there are different *distribution factors* that describe how different measures in grid operation effect the power flow on the lines. These distribution factors can be computed using the DC load flow calculation that we introduce in Subsection 3.1.3. In this subsection, we introduce three different kinds of distribution factors. The different distribution factors are used in Chapter 4 in our optimization model to define different kind of line outages. Further information on distribution factors can be found in [15]. In this section we use the following notations to denote the buses and lines in a power system that can be interpreted as an undirected graph. The set N describes the buses in the grid with $n \in N$ as a bus index. The set L describes the lines of the grid with elements $l, k \in L$. In addition the lines l and k can get identified with their start bus and end bus in the form $l = (l_1, l_2), k = (k_1, k_2) \in N^2$. The set L_{PST} describes the set of lines that are equipped with a phase-shifting transformer. For the line that is equipped with a PST we use the index $p \in L_{PST}$.

3.4.1 Power Transfer Distribution Factors

The most common used distribution factors are the so-called *power transfer distribution factors (PTDFs)*. The PTDFs can be interpreted as a matrix

$$PTDF \in \mathbb{R}^{L\times N},$$

with the value $PTDF_{l,n}$ describing the change of power flow in MW on line l when generating additional 1 MW of power at node n that is withdrawn from the grid at a reference bus that is fixed for all PTDFs.

3.4.2 Line Outage Distribution Factors

Using the PTDFs, we can describe the *line outage distribution factors (LODFs)* that can be interpreted as a matrix

$$LODF \in \mathbb{R}^{L\times L}.$$

The value $LODF_{l,k}$ is the dimensionless fraction of power that flows on line k and that gets redistributed to line l after an outage of line k. In fact we can compute $LODF_{l,k}$ as

$$LODF_{l,k} = \frac{PTDF_{l,k_1} - PTDF_{l,k_2}}{1 - (PTDF_{k,k_1} - PTDF_{k,k_2})},$$

with k_1 is the start bus of k and k_2 is the end bus of k. The numerator of this fraction corresponds to the power flow over line l of 1 MW of power that is generated at bus k_1 and withdrawn from the system at bus k_2. In the case of an outage of line k, this power flow is scaled by the reciprocal of $1 - (PTDF_{k,k_1} - PTDF_{k,k_2})$, the amount of power that flows from bus k_1 to k_2 but not on the line k. This relation holds since from the linearization of the load flow equations it follows that the power flow changes in a linear manner if all generations and loads are scaled by a constant factor.

3.4.3 Phase-shifting Distribution Factors

An easy way to describe the effects of PSTs, while using the linearized approach as presented in Subsection 3.3, is to look at the so-called *phase-shifting distribution factors (PSDFs)*. The PSDFs can be interpreted as matrix

$$PSDF \in \mathbb{R}^{L \times L_{PST}}.$$

The value $PSDF_{l,p}$ describes how the power flow on line l changes if the tap of the PST p is changed by 1°. The PSDFs can be computed directly from the grid data as described in [17] or by using the DC load flow calculation as described in Subsection 3.1.3.

4 Remedial Action Optimization in the Region Italy North

4.1 Software

To handle the data, the simulations, and the optimization we use different software packages and programming languages that we list here. The code for this master's thesis is written in the programming languages Julia and Python. The two most important software packages that we use are CPLEX which is a commercial solver for mathematical optimization problems and PyPSA an open source package for power system analysis that is developed using the programming language Python by the Karlsruhe Institute of Technology.

4.1.1 The Julia Programming Language

The linear optimization model that we develop and that gets further described in Section 4.3 is written in the *Julia programming language* [20]. The final code that we use for the computations in this thesis is implemented in Julia version 1.6.1.

4.1.2 The Python Programming Language

The data we use for our simulations is retrieved from the open optimization model PyPSA-Eur [21] as further described in Section 4.2. PyPSA-Eur relies on the software package PyPSA [22]. This software package is written using the *Python programming language* [23]. To make use of this data we have implemented an interface between Python and Julia to extract the data. The final code that we use for the simulations in this thesis is written using Python version 3.7.9.

4.1.3 CPLEX

CPLEX is a feature of the *IBM ILOG Optimization Studio* [24]. CPLEX is an optimization engine for solving problems that are expressed as mathematical programming models. For our computations we use IBM ILOG Optimization Studio V12.10.0. We do not use the CPLEX Optimization Studio directly but an interface integrated in the Julia module JuMP that allows us to write the optimization model in Julia and optimize it using CPLEX. We use the academic license for CPLEX that IBM provides at no charge to students and researchers.

4.1.4 PyPSA

The software toolbox *Pyhton for Power System Analysis (PyPSA)* can be used to simulate and optimize electrical power systems over multiple timestamps [22]. PyPSA includes models for different power system components as generators, buses, and line elements. Additionally, it contains predefined models for optimization of the operation and investment in power systems. Our main interest in PyPSA relies on the availability of the free open model PyPSA-Eur [21] of the European transmission system that we use in this thesis.

Figure 4.1: 512 bus grid

4.2 Data

4.2.1 Grid Data

To make the computations in this master's thesis reproducible, we only use open data. There are not many open data sources for the European transmission grid available. An open optimization model of the European transmission system is PyPSA-Eur [25]. The model of the grid contains 6001 lines. These lines are AC lines at and above a voltage level of 220 kV and all HVDC lines. Additionally, PyPSA-Eur contains an open database of conventional power plants, time series for electrical demand, and variable renewable generator availability. In [25] it is stated that the model is suitable for operational studies. The developers of PyPSA [22] do not recommend using the network with full resolution for simulations. Due to the way PyPSA-Eur builds the model from different data sources topological errors like the wrong assignment of loads and generators to network buses can arise. For this reason, it is recommended to cluster the network to a couple of hundred nodes to resolve local inconsistencies as described in [21]. We follow this recommendation and use a version of the grid that is preclustered and contains 512 buses and 888 lines. The clustered version of the PyPSA-Eur model that we use can be found on the web page [26]. Figure 4.1 shows the grid topology of the PyPSA-Eur grid with 512 buses that we use in our simulations.

4.2.2 Market Clearing

Remedial action optimization in the day-ahead framework requires as an input the predicted flows and the ordered generator dispatch for the day for each hour respectively. In Section 1.2.1, it is described that the market clearing is performed by an algorithm that maximizes social welfare, while matching bids and offers on the electricity market. Since we use open data to allow reproducibility, we perform a simplified market clearing using PyPSA. To reproduce the problems that arise in the market clearing algorithm, we only add capacities on cross zonal lines. The market clearing is performed by an optimization minimizing total costs for power generation while satisfying all power demands. Additionally, the PSTs are not considered in the market clearing and the tap positions are fixed to zero. The HVDCs get cooptimized with the generator dispatch. This optimization is included in PyPSA and called *linear optimal power flow*. All computations in this section are performed over the course of one day that is discretized into 24 equidistant timestamps, one for each hour of the day. This holds for the market clearing and the remedial action optimization. PyPSA-Eur includes data for loads for every hour of the year 2013. We choose a random day of this year that we use throughout our computations. The random day is the 25th of July. For other days we experience similar results to the ones we present in Subsection 4.4. We do not include different days in our comparison and focus on the effects of different model configurations for one day solely.

4.2.3 The Region Italy North

This master's thesis is written in cooperation with the Austrian Institute of Technology GmbH and the Austrian Power Grid GmbH which we introduce in Subsection 1.1. Therefore, the main focus of the simulations lies on Austria and its transmission system. The ROSC methodology we introduce in Section 1.2.2 establishes a coordination methodology for the region containing the countries Austria, Belgium, Croatia, Czechia, France, Germany, Hungary, the Netherlands, Luxembourg, Poland, Romania, Slovakia, Slovenia. To focus on the main aspects of the methodology, we apply the remedial action optimization to the smaller region *Italy North*. The region Italy North includes the countries Austria, France, Italy, Slovenia and Switzerland. The model still includes power flows, power generations and consumptions for the whole grid that is shown in Figure 4.1. In our model remedial actions are only available in the countries of the region Italy North.

4.2.4 PST Data

The PyPSA-Eur model does not contain data on the PSTs that are installed in the region Italy North. Therefore, we have to use data from external sources that we merge into the model. The PST Data we derive from the sources [27], [28] and [29] is listed in Table 4.1. For every PST its location, the country it is assigned to, its tap range, voltage level, and maximal continuous power is listed. Note that the PSTs in Padriciano and Divača are located very closely to each other. With the clustering that we use it is not sensible to differentiate between them and therefore we model them as one PST using the data of the PST in Divača which has a bigger tap range. Note that one has to transform the tap range from degree to radians to use it as a bound on the PST tap that we have introduce in Section 3.3. As the clustered grid does not contain all

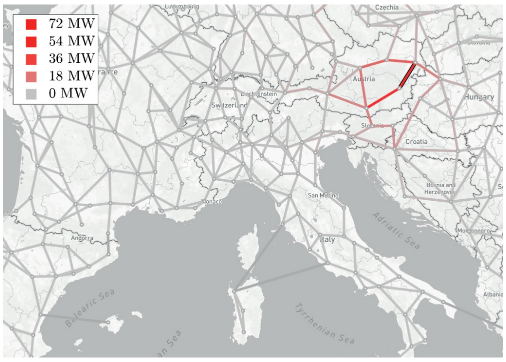
Country	Location	Tap range	Voltage level	Maximal continuous power
(AT)	Ernsthofen	±35°	220 kV	600 MVA
	Tauern	±35°	220 kV	600 MVA
	Ternitz	±35°	220 kV	600 MVA
	Lienz	±60°	220 kV	330 MVA
(IT)	Rondissone	±18°	400 kV	1630 MVA
	Camporosso	±54.9°	220 kV	450 MVA
	Foggia	±17.5°	400 kV	1800 MVA
	Padriciano	±31°	220 kV	370 MVA
(SI)	Divača	±40°	400 kV	1200 MVA

Table 4.1: PST data

lines, we have to assign the PSTs to lines in the 512 Bus grid. To later interpret the results, we investigate how the PSTs influence the power flows in the grid. In the same fashion as in [17], we visualize the PSDFs for the clustered grid in Figure 4.2. Each of the subfigures show the location of the PST on the line that has a black contour and the PSDF in MW for each line color coded with red for 72 MW and light grey for 0 MW. In all subfigures of Figure 4.2 one can clearly see that the PSTs have a local impact on the power flow in the grid.

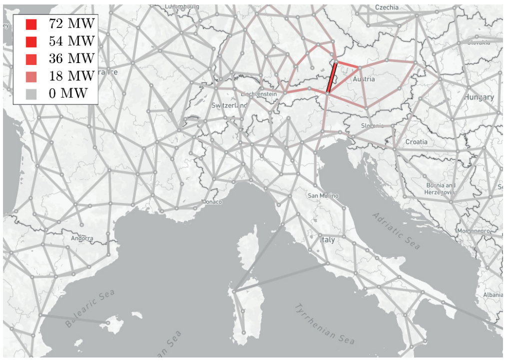
4.2.5 Contingencies

Contingencies are possible outages of transmission lines that we consider during grid operation. Considering all possible outages and their effects on every transmission line is not feasible because for every considered contingency our model grows in size. Taking into account all possible contingencies results in a huge increase in computation time and effort. Additionally, the outage of a line in the south of Italy does not effect the power flow on a line in Austria. Therefore, we define a contingency list that contains outages and monitored lines that we consider in the remedial action optimization.

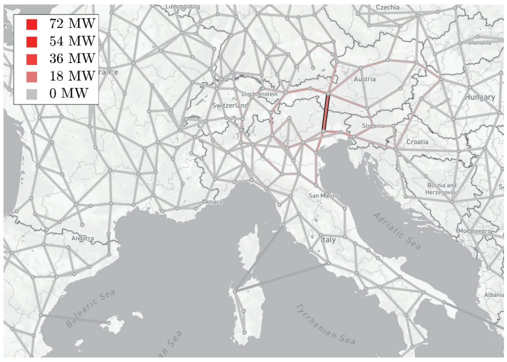


(a) PST in Ternitz (AT)

(b) PST in Ernsthofen (AT)

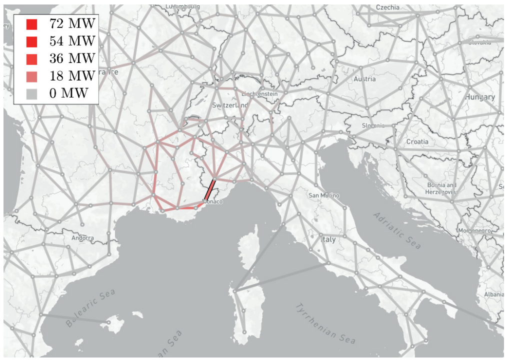


(c) PST in Tauern (AT)



(d) PST in Lienz (AT)

(e) PST in Rondissone (IT)

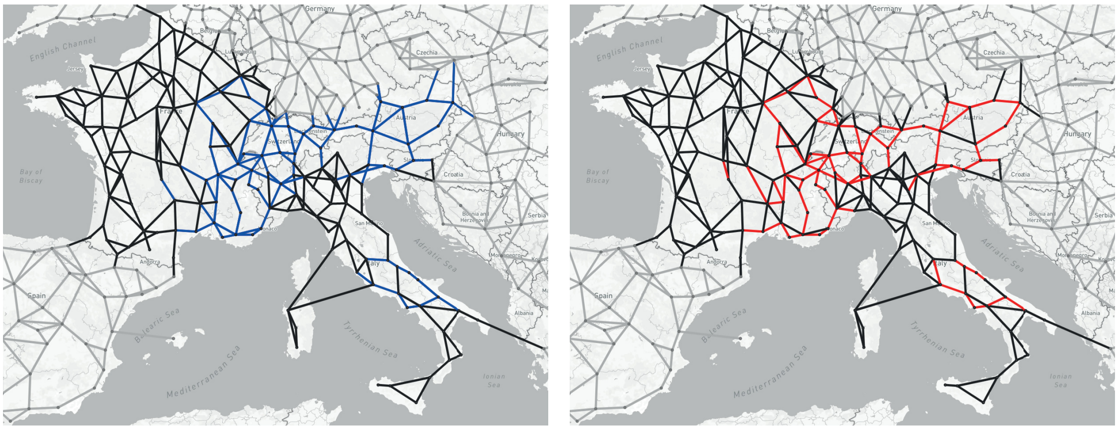


(f) PST in Camporosso (IT)

(g) PST in Foggia (IT)

(h) PSTs in Padriciano (IT), Divača (SI)

Figure 4.2: PSDFs for each PST



(a) all critical lines (b) all outage lines

Figure 4.3: Contingencies

To define the list of contingencies, we use the distribution factors from Section 3.4. Particularly, we use the line outage distribution factors that describe how the outage of a line effects the power flow on the other lines and the phase-shifting distribution factors that describe how the change in tap position of a phase-shifting transformer influences the power flow on each line.
We define the following contingencies using the LODF values:

$$C_{LODF} := \left\{ (l,k) \in L^2 \;\middle|\; \max_{k \in L} \left\{ |LODF_{l,k}| \right\} \geq .1 \text{ with } k = \arg\max_{k \in L} \left\{ |LODF_{l,k}| \right\} \right\}.$$

In other words, this means that for every line we only care for a possible outage if the worst case outage has a LODF with an absolute value greater or equal to 10% and then only include the worst case contingency.

In addition, we consider the following cross zonal contingencies that we define as

$$C_X := \left\{ (l,k) \in L^2 \mid |LODF_{l,k}| \geq .1 \text{ with } k \in L_X \right\},$$

where L_X is the set of lines $l \in L$, with the start bus and the end bus belonging to different countries. Therefore, C_X is the set of all cross-zonal contingencies, which have an LODF greater than or equal to 10% for the outage of a cross-zonal line.

Since the main focus of this master's thesis lies on the effect of PSTs in the remedial action optimization, we only consider contingencies where the power flow on the monitored line can be influenced by a PST. To this end, we use the PSDFs that we introduce in Subsection 4.2.4. We use them to define the set of contingencies that are influenced by a PST

$$C_{PST} := \left\{ (l,k) \in L^2 \;\middle|\; \max_{p \in L_{PST}} |PSDF_{l,p}| \geq 2 \right\}.$$

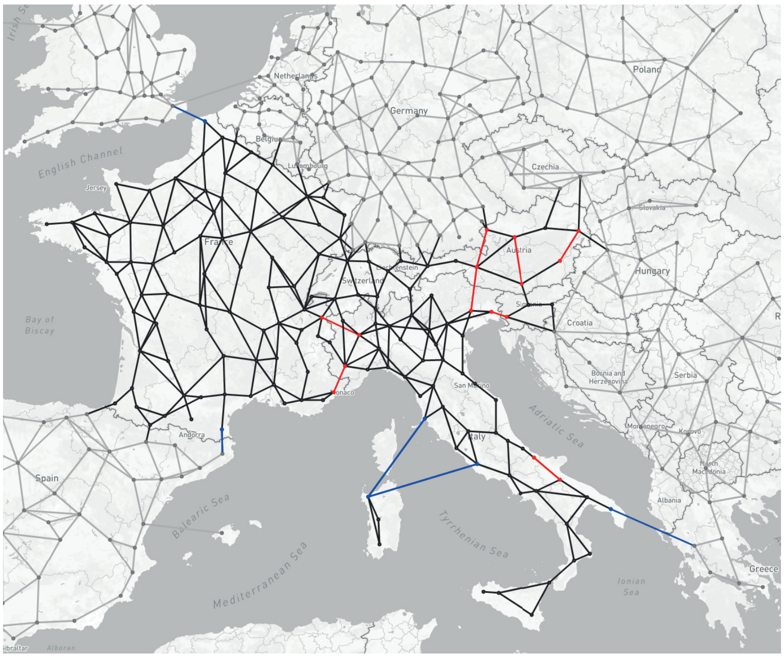


Figure 4.4: Grid of Italy North, lines in black, PSTs in red, HVDCs in blue

We only consider a contingency if there exists some PST that has a PSDF greater or equal to 2 for this line.

In total, the set of contingencies is defined as

$$C \coloneqq (C_{LODF} \cup C_X) \cap C_{PST}.$$

We consider all contingencies given by LODFs and all cross zonal contingencies that can get influenced by a PST to some extent. Figure 4.3 visualizes the set of contingencies. In Subfigure 4.3a all lines which get monitored for at least one outage are colored in blue. In Subfigure 4.3b all lines which fail at least once are colored in red. From the figure one can see that our choice of contingencies favors critical lines and outage lines that are close to a border in the region Italy North or are close to a PST. We use this set of contingencies in our model in Subsection 4.3.

4.3 The Optimization Model

In this section, we describe the linear model for remedial action optimization that we develop in this master's thesis. The main inspiration for the equations is taken from the paper on PyPSA [22]. In this source linear power flow equations are formulated and the optimization of operation and investment is described. Another important inspiration is the open source toolbox *FARAO* which stands for *Fully Autonomous Remedial Actions*

Optimisation [30]. FARAO aims at providing a complete toolbox for remedial action optimization. Since FARAO is still under active development, we develop an own model to investigate the effects of non-costly remedial actions in the remedial action optimization.

4.3.1 Model Inputs and Notation

In this subsection, we list the mathematical notations we use to formulate our optimization model. All notations are listed in Table 4.2. The table includes for all variables their names, a brief description, and a reference to a previous subsection with a description of the variable. The variable names are partially aligned with the notations in [22].

4.3.2 The main Constraints

In this subsection, we describe the main equations of our optimization model. We start with the linear load flow equations that we derive in Subsection 3.1.2. For every bus in the grid $n \in N$ and for every timestamp $t \in T$ we introduce the phase angles $\theta_{n,t} \in \mathbb{R}$. Then for every line $l \in L_{line}$ using the line reactance x_l, we get the following relation for the linear power flow $f_{l,t}$ on the line l at timestamp t

$$f_{l,t} = \frac{\theta_{n,t} - \theta_{m,t}}{x_l} \quad \forall l = (n,m) \in L_{line},\ t \in T,$$

with the line $l \in L_{line}$ identified with the ordered tuple consisting of its start bus n and end bus m in the form $(n,m) \in N^2$. We also include constraints for the maximal allowable power flow on each line. As described in Subsection 4.2.2, it might happen that the predicted market flows do not respect these constraints. These violations have to be resolved using remedial actions. Therefore, we add the constraints

$$-f_l^{max} \leq f_{l,t} \leq f_l^{max} \quad \forall l \in L_{ITN},\ t \in T, \tag{4.1}$$

and

$$-\max(f_l^{max}, |f_{l,t}^{market}|) \leq f_{l,t} \leq \max(f_l^{max}, |f_{l,t}^{market}|) \quad \forall l \in L_{outsiede},\ t \in T, \tag{4.2}$$

with f_l^{max} the line limit and $f_{l,t}^{market}$ the power flow that gets predicted by the market clearing. The constraint (4.1) describes that within the region Italy North the remedial action optimization shall produce a state where no overloads appear. The constraint (4.2) describes that outside of Italy North the coordination will not worsen predicted overloads. We use these weaker constraints outside of Italy North since we are only able to use remedial actions within the region Italy North and can not guarantee to resolve overloads that are far distant from this region.

For the PSTs in accordance with Subsection 3.3, we also need to consider the tap setting $\alpha_{l,t}$ for $l \in L_{PST}$ and $t \in T$ and the reactance x_l^{PST} to derive the formulation

$$f_{l,t} = \frac{\theta_{n,t} - \theta_{m,t} + \alpha_{l,t}^{PST}}{x_l + x_l^{PST}} \quad \forall l = (n,m) \in L_{PST},\ t \in T.$$

Variable	Description	Reference
$n, m \in N$	bus labels and set of bus labels	Subsection 4.2.1
$l, k \in L$	line labels and set of line labels	Subsection 4.2.1
$L_{line} \subset L$	subset of passive lines	Subsection 4.2.1
$L_{PST} \subset L$	subset of lines with PST	Subsection 4.2.4
$L_{HVDC} \subset L$	subsect of HVDC lines	Subsection 4.2.1
$L_{ITN} \subset L$	subset of lines that belong to Italy North	Subsection 4.2.1
$L_{outside} \subset L$	subset of lines that do not belong to Italy North	Subsection 4.2.1
$t \in T$	timestamp and set of timestamps	Subsection 4.2.1
$\theta_{n,t}$	bus voltage angle at bus n at time t	Subsection 3.1.3
$f_{l,t}$	power flow on line l at time t	Subsection 3.1.3
x_l	reactance of line l	Subsection 3.1.3
x_l^{PST}	additional reactance of PST on line l	Section 3.3
f_l^{max}	line capacity on line l	Subsection 3.1.3
$\alpha_{l,t}$	PST tap setting	Section 3.3
α_l^{lower}	lower tap bound	Section 3.3
α_l^{upper}	upper tap bound	Section 3.3
$\alpha_{l,t}$	PST tap setting	Section 3.3
$LODF_{l,k}$	line outage distribution factor for lines l and k	Subsection 3.4.2
$c \in C \subset L^2$	contingency labels and set of contingencies	Subsection 4.2.5
$s \in S_n$	generator label and set of generators attached to bus n	Subsection 4.2.2
$g_{s,t}^{market}$	dispatch of generator s at timestamp t	Subsection 4.2.2
g_s^{max}	capacity of generator s	Subsection 4.2.2
$r_{s,t}$	redispatch of generator s at timestamp t	Subsection 4.2.2
c_s	costs of producing 1 MW at generator s	Subsection 4.2.2
$p_{n,t}^{market}$	nodal net position of bus n at timestamp t after market clearing	Subsection 4.2.2
$f_{l,t}^{market}$	power flow on line l at timestamp t after market clearing	Subsection 4.2.2

Table 4.2: Notations

We identify the line l that is equipped with a PST with the ordered tuple of its start bus and end bus $(n,m) \in N^2$. For the tap setting, we consider the following constraints

$$\alpha_l^{lower} \leq \alpha_{l,t} \leq \alpha_l^{upper} \quad \forall l \in L_{PST},\ t \in T,$$

that we derive from the data that we summarize in Subsection 4.2.4. We have to be aware that while the tap setting for a real PST is discrete, we stick to continuous values in order to formulate the remedial action optimization as a linear optimization problem as defined in Section 2.1. We also add constraints on the power flow for each timestamp as

$$-f_l^{max} \leq f_{l,t} \leq f_l^{max} \quad \forall l \in L_{PST},\ t \in T.$$

The power flow on the HVDC lines $f_{l,t}$ for $l \in L_{HVDC}$ and $t \in T$ is independent from the voltage angles in the grid and allows for full power control as described in Subsection 3.2. Therefore, we only have to restrict the flow on the HVDC line for all timestamps using the line capacities F_l^{max} to get

$$-f_l^{max} \leq f_{l,t} \leq f_l^{max} \quad \forall l \in L_{HVDC},\ t \in T.$$

After modeling the non-costly remedial actions, we now take a look at the generator redispatch. For every bus $n \in N$ the set S_n is an index set with one index for each generator that is attached to the bus n. For each generator $s \in S_n$ and timestamp $t \in T$, we define the redispatch as $r_{s,t}$. The available redispatch depends on the ordered dispatch for a timestamp t and is constrained as

$$r_{s,t}^{lower} \leq r_{s,t} \leq r_{s,t}^{upper} \quad \forall n \in N,\ s \in S_n,$$

with $r_{s,t}^{lower} := -g_{s,t}^{market}$ and $r_{s,t}^{upper} := g_s^{max} - g_{s,t}^{market}$. Here $g_{s,t}^{market}$ is the ordered dispatch for generator s at timestamp t and g_s^{max} is the maximal capacity of the generator. The constraint on the redispatch ensures that every generator after applied redispatch generates power within its given bounds. We also include ramp limits that constrain how much the power generation of a generator can be changed from one timestamp to the next. They are included in the model in the form

$$\begin{gathered} -0.2 \cdot g_s^{max} \leq (g_{s,t_k} + r_{s,t_k}) - \left(g_{s,t_{k+1}} + r_{s,t_{k+1}}\right) \leq 0.2 \cdot g_s^{max} \\ \forall n \in N,\ s \in S_n,\ k \in \{1, \dots, |T|-1\}, \end{gathered}$$

using the ordered notation $T = \{t_k \mid k = 1, \dots, |T|\}$. The dispatch of a generator taking redispatch into account shall respect the ramp limit of 20% of the generator capacity g_s^{max}. Now, we are able to combine the equations for the different flow types $f_{l,t}$, the market clearing net positions $p_{n,t}^{market}$, and the generator redispatch $r_{s,t}$ using the *Kirchhoff current law (KCL)* in the form of

$$p_{n,t}^{market} + \sum_{s \in S_n} r_{s,t} + \sum_{l=(m,n)} f_{l,t} - \sum_{l=(n,m)} f_{l,t} = 0 \quad \forall n \in N,\ t \in T.$$

The KCL states that for every bus $n \in N$ the amount of power generated at that bus minus the amount of power consumed at that bus plus the inflowing power minus the outflowing power equals zero. In the grid power can only be created by generation and destroyed by consumption. From there we can directly conclude that

$$\sum_{n \in N} \sum_{s \in S_n} r_{s,t} = 0 \quad \forall t \in T.$$

This identity can easily be derived by summing the KCL constraints over all $n \in N$ and noting that the flows cancel away because by adding the flows for all buses every flow get once added when summing with respect to its start bus and once subtracted when summing with respect to its end bus and the fact that the market clearing results in nodal net positions that satisfy $\sum_{n \in N} p_{n,t}^{market} = 0$ for all $t \in T$. Therefore, we can conclude that the total redispatch is balanced at every timestamp.
In the security analysis, we consider different contingencies as described in Subsection 4.2.5. A contingency $c \in C$ consists of an ordered tuple $c = (l, k) \in L^2$, where $l \in L$ is a critical line and $k \in L$ is the line that experiences an outage. In the case of an outage of line k a portion of the flow $f_{k,t}$ gets redistributed to the line l. The post outage flow on line l is then $f_{l,t} + LODF_{l,k} f_{k,t}$. In general, we cannot hope that for every contingency and every timestamp we can prevent overloads. Therefore, we cannot use the direct constraint

$$-f_l^{max} \leq f_{l,t} + LODF_{l,k} f_{k,t} \leq f_l^{max} \quad \forall c \in C,\ t \in T, \tag{4.3}$$

while guaranteeing feasibility of the model. Instead for each contingency c and each timestamp t, we introduce the variable

$$O_{c,t} \geq 0 \quad \forall c \in C,\ t \in T. \tag{4.4}$$

and add it to the right-hand side and subtract it from the left-hand side of (4.3) to arrive at

$$-\left(f_l^{max} + O_{c,t}\right) \leq f_{l,t} + LODF_{l,k} f_{k,t} \leq f_l^{max} + O_{c,t} \ \forall c \in C,\ t \in T. \tag{4.5}$$

A mathematical optimization model as defined in Section 2.1 needs an objective function. In the following subsection, we define three different model configurations with different objective functions to further investigate the effect of non-costly remedial actions in the remedial action optimization.

4.3.3 Model Configurations

To investigate the effects of PTSs and HVDCs in the remedial action optimization, we consider three different model configurations with different objective functions and different additional constraints. Using these model configurations, we compare how the use of non-costly remedial actions (PSTs and HVDCs) benefits remedial action optimization in comparison to the pure use of generator redispatch. Before we formulate the different objective functions, we give an additional remark.

Remark 4.3.1.
We use sums of absolute values when we define the different objective functions. The absolute value is obviously not a linear function. Nevertheless, the minimization of absolute values can be expressed by a linear program. Because we do not want to apply this reformulation to every objective function, we explain the procedure once in this remark.
Consider the objective function

$$f : \mathbb{R}^n \to \mathbb{R},\ x \mapsto \sum_{i=1}^{n} c_i |x_i| ,$$

for a fixed $c \in \mathbb{R}^n$, with $c \geq 0$. We want to solve the nonlinear optimization problem

$$\begin{array}{lr} \min & f(x) \\ s.t. & x \in \mathbb{R}^n, \\ & g(x) \leq 0, \\ & h(x) = 0, \end{array}$$

with affine linear constraints $g : \mathbb{R}^n \to \mathbb{R}^m$ and $h : \mathbb{R}^n \to \mathbb{R}^p$. We want to solve this problem using algorithms that are suitable for linear optimization. Therefore, we use the following reformulation. First, we introduce the variable

$$x^+ \in \mathbb{R}^n,$$

and add the two componentwise constraints

$$\begin{aligned} x^+ &\geq x \quad and \\ x^+ &\geq -x. \end{aligned}$$

Next, we define the linear objective function

$$f^+(x^+) = c^T x^+.$$

Then, we define the linear optimization problem

$$\begin{array}{lr} \min & f^+(x^+) \\ s.t. & (x, x^+) \in \mathbb{R}^{2n}, \\ & g(x) \leq 0, \\ & h(x) = 0, \\ & x^+ \geq\ x, \\ & x^+ \geq -x. \end{array}$$

The two optimization problems are equivalent in the following sense.

- *The feasible set of one problem is non-empty if and only if the feasible set of the other problem is non-empty.*

- *One problem has an optimal solution if and only if the other problem has an optimal solution.*
- *If x is an optimal solution to the first problem, then defining $x^+ \in \mathbb{R}^n$ with $x_i^+ = |x_i|$ for $i = 1, \dots, n$ yields an optimal solution (x, x^+) to the second problem.*
- *If (x, x^+) is an optimal solution to the second problem, then x is an optimal solution to the first problem .*

We apply this technique in the remainder of this subsection to define the objective functions for the remedial action optimization model.

RAO Model Configuration 1

Using the first objective function, we aim for two goals. On the one hand side, we want to minimize the total overloads that get expressed as $\sum_{c \in C,\ t \in T} O_{c,t}$. On the other hand side, we also want to minimize the total costs that emerge due to redispatch. They are equal to $\sum_{n=1}^{N} \sum_{s \in S_n} |r_s|\, c_s$. In the sum $|r_s|$ is the absolute amount of redispatch at generator s and c_s is the costs for producing 1 MW of electrical energy with generator s. Both of these different objectives get combined in one objective function using the parameterized function

$$\sum_{n \in N} \sum_{s \in S_n} |r_s|\, c_s + c^{cong} \sum_{c \in C,\ t \in T} O_{c,t},$$

with $c^{cong} > 0$ as a weighting factor. In our simulations, we choose $c^{cong} = 150$. In RAO model configuration 1, we fix the PST tap settings and HVDC flows in order to understand their effects on the remedial action optimization. Therefore, we add two additional constraints. First, we fix the HVDCs to the ordered flows that get computed in the market clearing step

$$f_{l,t} = f_{l,t}^{market} \quad \forall l \in L_{HVDC},\ t \in T, \tag{4.6}$$

and additionally, we fix the PST tap settings using

$$\alpha_{l,t} = 0 \quad \forall l \in L_{PST},\ t \in T. \tag{4.7}$$

This way just as in the market clearing, we treat the PSTs like they are nonexistent and just passive lines.

RAO Model Configuration 2

In RAO model configuration 2, we use the same objective function as in RAO model configuration 1 that reads as

$$\sum_{n\in N}\sum_{s\in S_n} |r_s|\, c_s + c^{cong} \sum_{c\in C,\ t\in T} O_{c,t}, \tag{4.8}$$

with $c^{cong} = 150$. In RAO model configuration 2, we do not include the additional constraints (4.6) and (4.7) and allow control of the non-costly remedial actions.

RAO Model Configuration 3

RAO model configuration 2 allows full control of the PSTs and HVDCs. In simulations this leads to PST tap settings that vary from hour to hour. Hourly changes of tap positions can be impractical as they lead to higher wear and tear of the PSTs. Therefore, in the objective function of RAO model configuration 3, we add additional penalty terms for the usage of PSTs and HVDCs. The penalty term on the PSTs is

$$\sum_{l\in L_{PST}} \sum_{k=1}^{|T|-1} |\alpha_{l,t_k} - \alpha_{l,t_{k+1}}|, \tag{4.9}$$

to punish hourly changes in the tap position. When controlling the HVDCs, we do not want unnecessary changes from the ordered operation that gets computed during the market clearing. Therefore, we add the penalty term

$$\sum_{l\in L_{HVDC}, t\in T} |f_{l,t} - f_{l,t}^{market}|. \tag{4.10}$$

We introduce the two parameters $c^{PST} > 0$ and $c^{HVDC} > 0$ and combine the objective function (4.8) with the PST penalty term (4.9) and the HVDC penalty term (4.10) to arrive at

$$\begin{gathered}\sum_{n\in N}\sum_{s\in S_n} |r_s|\, c_{n,s} + c^{cong} \sum_{c\in C,\ t\in T} O_{c,t} + \\ c^{PST} \sum_{l\in L_{PST}} \sum_{k=1}^{|T|-1} |\alpha_{l,t_k} - \alpha_{l,t_{k+1}}| + c^{HVDC} \sum_{l\in L_{HVDC}, t\in T} |f_{l,t} - f_{l,t}^{market}|.\end{gathered} \tag{4.11}$$

In the simulations, we choose the parameters $c^{cong} = 150$, $c^{PST} = 500$, and $c^{HVDC} = 10$.

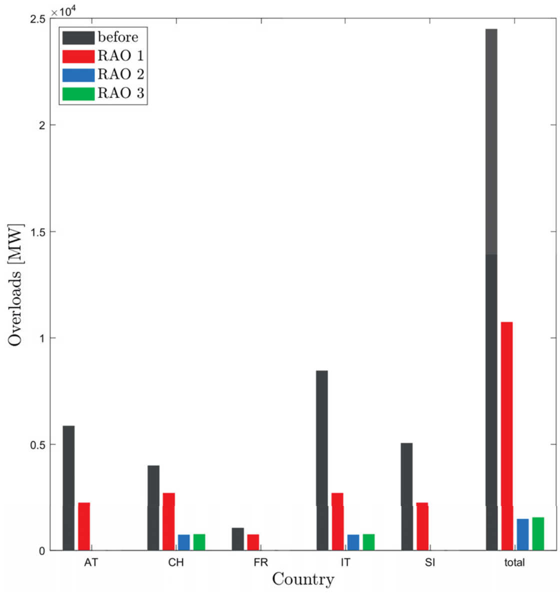


Figure 4.5: Total overloads

4.4 Results

For each of the RAO model configurations from the previous section we use CPLEX to optimize the model with the interior-point method that gets described in Chapter 2. We then compare the results in three parts. First, we look for every RAO model configuration at the total overloads. Then, we compare the incurred costs and the applied redispatch. After that, we compare the hourly PST operation in the course of the day. To be concise in the remainder of this section instead of "RAO model configuration 1/2/3" we only write "RAO 1/2/3".

4.4.1 Overloads

In the previous subsection, we describe the overloads $O_{c,t}$ using the inequalities (4.4) and (4.5). In each RAO model configuration, a component we want to minimize is the sum of overloads $\sum_{c\in C,\ t\in T} O_{c,t}$. We compare for each RAO model configuration to what extend the overloads get minimized. To get a more detailed insight into the overloads, we investigate the overloads per country. Each bus $n \in N$ is assigned to a country. When comparing the overloads per country we assign the overload $O_{c,t}$ with contingency $c = (l, k)$ and monitored line $l = (n, m)$ in equal parts to the countries that the start bus n and the end bus m are assigned to. Figure 4.5 is a bar chart showing the total overloads per country and the total overloads for Italy North in MW before the remedial action optimization and after the remedial action optimization for each of the three RAO model configurations. The overloads in MW are on the y-axis. There are four bars for each country (AT, CH, FR, IT, SI) and for the whole region Italy North (total) that are listed on the x-axis. The black bar shows the overloads that emerge due to contingencies

Country	before	RAO 1	RAO 2	RAO 3
AT	5,886.66	2,273.07 (38.6%)	0.00 (0.0%)	20.16 (0.3%)
CH	4,024.06	2,728.20 (67.8%)	740.99 (18.4%)	765.16 (19,0%)
FR	1,059.16	758.70 (71.6%)	0.00 (0.0%)	0.00 (0.0%)
IT	8,467.40	2,728.20 (32.2%)	740.99 (8.8%)	765.16 (9.0%)
SI	5,074.33	2,271.56 (44.8%)	0.00 (0.0%)	0.00 (0.0%)
total	24,511.62	10,759.71 (43.9%)	1,481.97 (6.1%)	1,550.49 (6.3%)

Table 4.3: Total overloads

in the situation where remedial actions are not activated. The red bar stands for RAO 1 where only costly remedial actions get optimized. The blue bar stands for RAO 2 where costly remedial actions and non-costly remedial actions get optimized. The green bar stands for RAO 3 where costly and non-costly remedial actions get optimized and we impose additional penalty terms on the PST tap settings and the ordered HVDCs flows. Table 4.3 lists the total overloads per country and the total overloads for Italy North in MW before the remedial action optimization and after the remedial action optimization for each of the three RAO model configurations as numerical values. Additionally, the table includes for each model configuration the overloads after activation of remedial actions relative to the overloads before activation of remedial actions in percentage to one decimal place. From Figure 4.5 and Table 4.3 we can draw the following conclusions. From the bar chart one can clearly see that each of the RAO model configurations reduces the overloads for each country respectively and for the whole region Italy North. Using only costly remedial the overloads per country can be reduced between 32.2% (IT) and 71.1% (FR) of the total overload value before activation of remedial actions. The total overloads for Italy North using only costly remedial action get reduced from 24,511.62

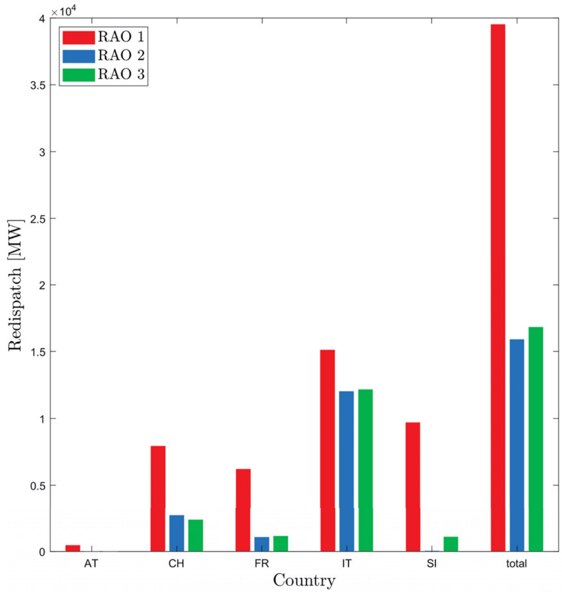


Figure 4.6: Optimized redispatch

MW to 10,759.71 MW. This is a reduction to 43.9% of the original value. The overload results for RAO 2 and 3 show the power of the non-costly remedial actions. The model of the grid contains only 8 PSTs and 5 HVDCs. These additional coordination measures reduce the total overloads even further. For the RAO 2 the overloads in Austria (AT), France (FR) and Slovenia (SI) get reduced to zero. Only in Switzerland (CH) and Italy (IT) there remain overloads of 740.99 MW for both countries which in total correspond to overloads of 1481.97 MW for the whole region of Italy North (total). Using RAO 2, the total overloads get reduced to 6.1% of the value before the activation of remedial actions. RAO 3 performs only a little worse than RAO 2. In Table 4.3 one can see that for no country the overloads using RAO 3 are increased by more than 25 MW for each country and in total by less than 70 MW compared to RAO 2. This increase is only minor as it only accounts for 0.2% of the overloads before activation of remedial actions for the whole region Italy North. In conclusion, we see that all RAO model configurations significantly decrease the overloads and that the usage of non-costly remedial actions provides significant additional improvement.

4.4.2 Redispatch and Costs

In this subsection, we compare the total redispatch and costs for each of the RAO model configurations starting with the redispatch. To get a more detailed overview we compare the redispatch and costs for every country respectively. Each bus $n \in N$ in the grid is assigned to a country. Then, for each country $Z \in \{AT, CH, FR, IT, SI\}$ we denote by $N_Z \subset N$ the buses that are located in the country Z and compute the total redispatch

Country	RAO 1	RAO 2	RAO 3
AT	474.67	0.00 (0.0%)	0.00 (0.0%)
CH	7,957.44	2,711.30 (34.1%)	2,386.73 (30.0%)
FR	6,238.01	1,088.75 (17.5%)	1,162.63 (18.5%)
IT	15,148.75	12,061.25 (79.6%)	12,206.47 (80.6%)
SI	9,729.29	69.95 (0.7%)	1,103.69 (11.3%)
total	39,548.15	15,931.24 (40.3%)	16,859.52 (42.6%)

Table 4.4: Optimized total redispatch in MW

for a country as $\sum_{n \in N_Z} \sum_{s \in S_n} |r_s|$. Additionally, we consider the total redispatch for the whole region Italy North that is computed by summing $\sum_{n \in N} \sum_{s \in S_n} |r_s|$. Figure 4.6 shows a bar chart that contains three bars for each country (AT, CH, FR, IT, SI) and the whole region Italy North (total) and shows the absolute redispatch. The bars are color coded with red corresponding to RAO 1, blue corresponding to RAO 2 and green corresponding to RAO 3. Additionally, in Table 4.4 the same data is listed as numerical values and it furhter contains for RAO 2 and 3 percentage values comparing it relative to RAO 1. From Figure 4.6 one can see that RAO 2 and 3 result in remedial action choices that use less redispatch than RAO 1. This holds for all countries on their own and for the region Italy North in total. This result is expected since in RAO 1 non-costly remedial actions can not be activated in order to resolve overloads. It is still noteworthy that the few PSTs and HVDCs account for such a significant decrease in redispatch. From Table 4.4 one can see that RAO 1 uses in total 39,548.15 MW of redispath while RAO 2 uses 15,931.24 MW of redispatch and RAO 3 uses 16,859.52 MW of redispatch. For RAO 2 and 3 this is a reduction of redispatch of ca. 60% in comparison to RAO 1. When comparing RAO 2 and 3, one notes that for all countries except Austria (AT) and Switzerland (CH) and for the whole region of Italy North there is an increase in redispatch from configuration 2 to 3. In Austria there is no use for redispatch at all. A possible reason for this is that four of eight PSTs are located in Austria and can be used very efficient to control load flows. In Switzerland the amount of redispatch decreases by ca. 325 MW when the non-costly remedial actions get restricted by additional penalty terms in the objective functions. France (FR) and Italy (IT) experience only slight increases of redispatch in comparison to the redispatch that is used in RAO 1. Slovenia (SI) has to invest the most in additional redispatch when the non-costly remedial actions get restricted by penalty terms and has to increase redispatch by ca. 1000 MW.

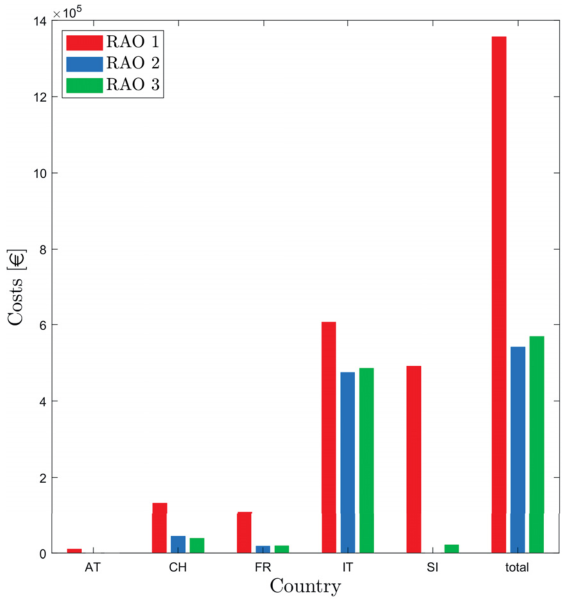


Figure 4.7: Costs for optimization

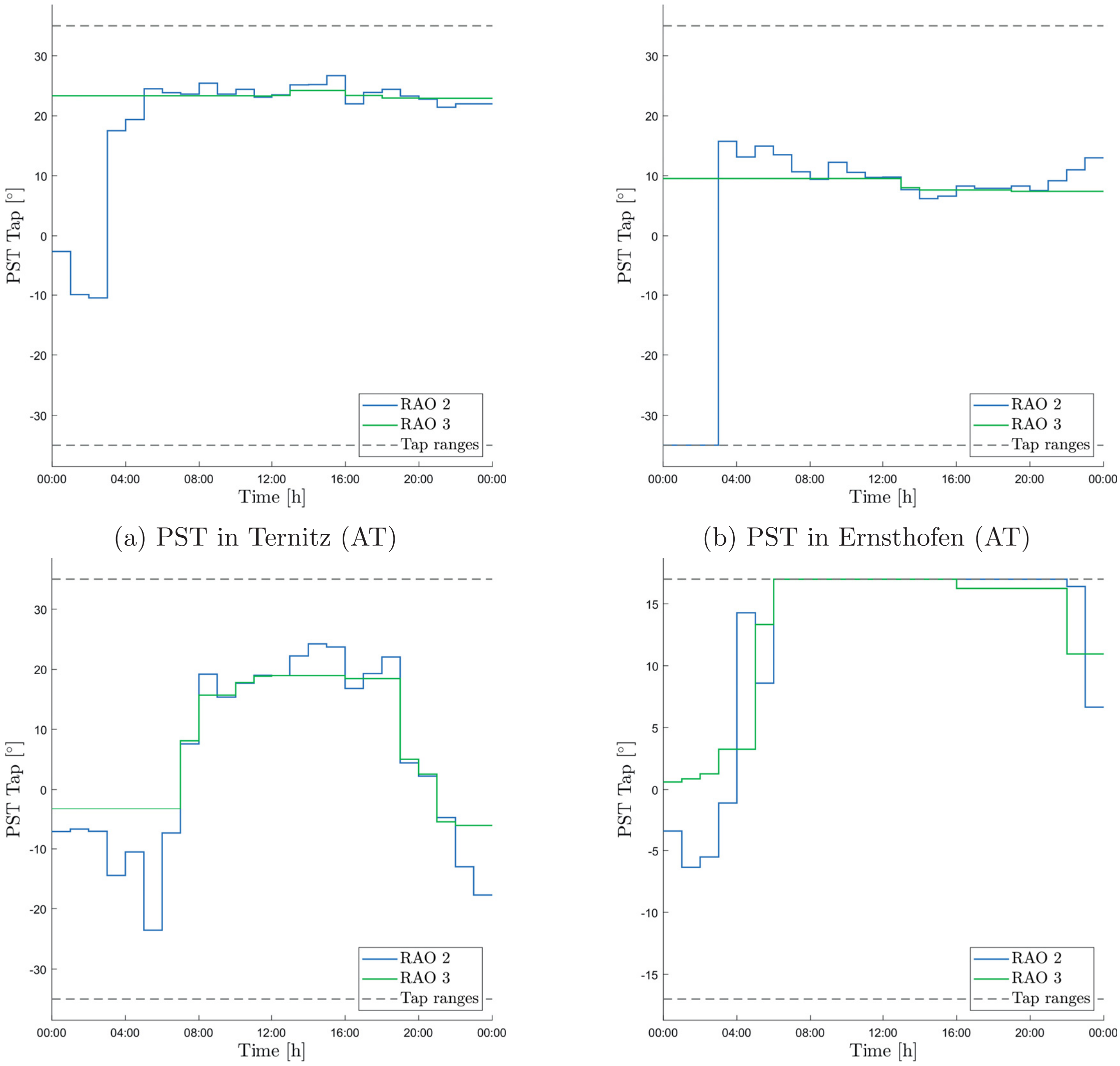
In the following, we compare the total costs that are incurred due to redispatch. The computation of the costs follows similar to the computation of redispatch. For each generator the amount of redispatch in MW gets multiplied with the costs for generating 1 MW at the respective generator. So for each country $Z \in \{AT, CH, FR, IT, SI\}$, we compute the costs per country as $\sum_{n \in N_Z} \sum_{s \in S_n} |r_s| \, c_s$ and the costs for the whole region Italy North as $\sum_{n \in N} \sum_{s \in S_n} |r_s| \, c_s$. The formulas for the redispatch and the costs are very similar and only differ by the weighting factors c_s. Therefore, the results are strongly correlated. Figure 4.7 shows a bar chart presenting the total costs per country and the total costs for Italy North in EUR before the remedial action optimization and after the remedial action optimization for each of the three RAO model configurations. The costs in EUR are on the y-axis. There are three bars for each country (AT, CH, FR, IT, SI) and for the whole region Italy North (total) that are listed on the x-axis. The bars for the arising costs are colored in red for RAO model configuration 1, blue for RAO model configuration 2 and green for RAO 3. Additionally, Table 4.5 shows the same data listed as numerical values and further for RAO 2 and 3 the table shows the costs in proportion to the costs of RAO 1 in percentage to one decimal place. We can see for all countries and when looking at the whole region Italy North that RAO 1 results in the highest costs. This is expected since with RAO 2 and 3 we allow the use of non-costly remedial actions. Nevertheless, it is remarkable that with the use of only eight PSTs and five HVDCs we are able to cut the total costs by 60% using RAO 2 and by 58% using RAO 3 in comparison to RAO 1 as we can see in Table 4.5. Again, we see that in Austria there are very little costs in comparison to the total value using RAO 1 and no costs using RAO 2 and 3. For Switzerland and France the costs using RAO 2 are slightly higher than the costs using RAO 3 contrary to Italy and Slovenia where it is the other way around. A possible reason for this is that Italy and Slovenia suffer more from the additional regularization of the PST tap settings that are imposed by the objective function of RAO 3. The difference in costs between RAO 2 and 3 is for all countries and the whole region Italy North less than 5% of the costs that occur using RAO 1. When looking only at the costs RAO 2 is the best outperforming RAO 3 only slightly. Both RAO 2 and 3 are superior to RAO 1.

Country	RAO 1	RAO 2	RAO 3
AT	11,441.13	0.00 (0.0%)	0.00 (0.0%)
CH	134,497.33	45,826.59 (34.1%)	40,340.67 (30.0%)
FR	110,784.33	19,562.87 (17.7%)	20,403.93 (18.4%)
IT	608,182.60	476,337.40 (78.3%)	486,902.81 (80.1%)
SI	493,098.86	943.40 (0.2%)	22,436.59 (4.6%)
total	1,358,004.27	542,670.26 (40.0%)	570,084.00 (42.0%)

Table 4.5: Total costs in EUR

4.4.3 PST Tap Settings

In the previous subsection, we compare the overloads, costs and the redispatch for the three different RAO model configurations. Overall, RAO 2 performes the best with respect to the overloads, costs and redispatch. The difference between RAO 2 and 3 lies in the different objective functions. In RAO 3 the focus shifts from the pure minimization of costs and overloads to a more regularized approach with additional penalty terms. The main motivation behind the additional penalty terms is to stop the optimization



(a) PST in Ternitz (AT)

(b) PST in Ernsthofen (AT)

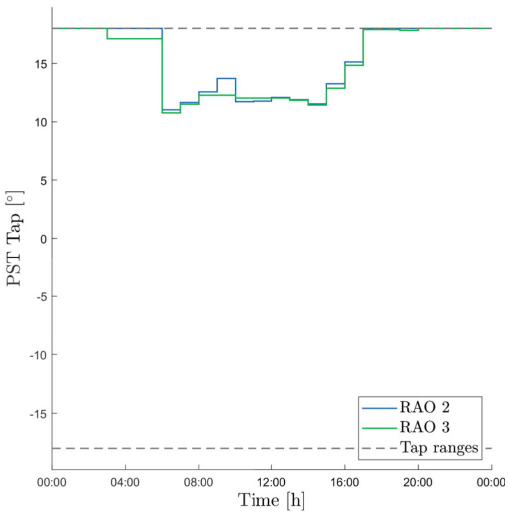
(c) PST in Tauern (AT)

(d) PST in Lienz (AT)

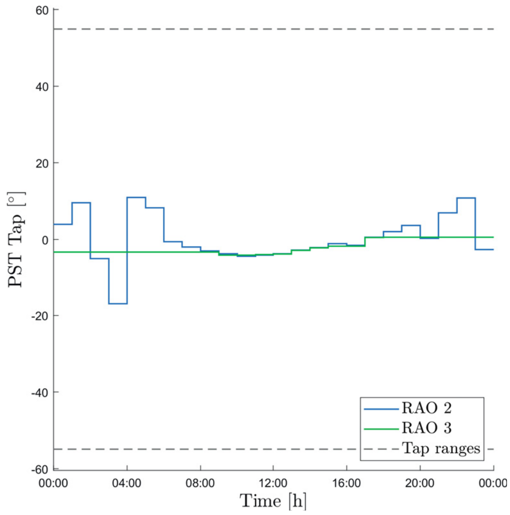
Figure 4.8: Optimized PST taps part 1

from producing results with PST tap positions that change every hour. To check if this regularization is successful, we investigate the times series of tap positions $(\alpha_{l,t})_{t \in T}$ for each PST $l \in L_{PST}$. In Figure 4.8 and 4.9 there are subfigures for each PST in the grid model. Each subfigure contains a line plot with the time in hours on the x-axis and the PST tap in degrees on the y-axis. Each plot contains the lower and upper tap range as a dashed grey line. The time series of the optimized tap positions are plotted in blue for RAO 2 and in green for RAO 3. We compare the resulting time series of all plots starting with Figure 4.8 that contains the four Austrian PSTs in Ternitz, Ernsthofen, Tauern and Lienz. In Subfigure 4.8a and 4.8b the PST tap series for RAO 3 is almost constant and over most of the time scale close to the time series for RAO 2 with exception of the first three hours. In Subfigure 4.8c, the time series of both model configurations align with the time series for RAO 3 containing to constants periods from 0:00 - 7:00 h and from 11:00 - 19:00 h. In Subfigure 4.8d, the time series look similar again. The time series for RAO configuration 3 has less oscillations but the number of hourly tap changes is comparable with RAO 2 having for this tap a longer constant period from 6:00 - 22:00 h.

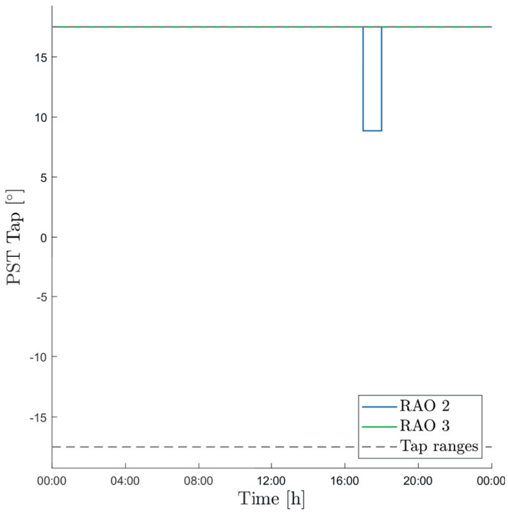
Figure 4.9 contains the Italian and Slovenian PSTs in Rondissione, Camporosso, Foggia and the two PSTs in Padriciano and Divača that are treated as one PST. In Subfigure 4.9a the time series for both RAO model configurations are again closely aligned. While RAO 3 produces an additional tap change at 3:00 and 4:00 h it produces a more constant tap positions from 8:00 - 13:00 h compared to RAO 2. In Subfigure 4.9b once again RAO 3 results in tap positions with less oscillations that still align with the ones from RAO 2. While the taps from RAO 2 have oscillations from 0:00 - 9:00 h and 17:00 - 0:00 h the taps from RAO 3 remain constant in this area. In the time frame 9:00 - 17:00 h the time series are almost identical. In Subfigures 4.9c and 4.9d all time series are almost constant with the time series of RAO 3 having even less hourly tap changes. In Subfigure 4.9c for RAO 2 there is only one tap position from 17:00 - 18:00 h that differs from the upper tap range while the time series for RAO 3 remains constant over the whole time range. In Subfigure 4.9d the time series for both model configurations remain constant from 3:00 - 0:00 h. While RAO 2 produces three tap changes between 0:00 - 3:00 h the time series produced by RAO 3 remains constant in this time range on a smaller tap position. All in all, we see that the regularization approach imposed by the additional penalty terms in the objective function of RAO 3 are successful. Overall, we see less hourly tap changes and longer time periods with constant tap positions while still maintaining tap time series that are similar to the ones that result from RAO 2. Choosing RAO 3 we have less wear and tear for each PST while maintaining qualitatively similar PST settings.



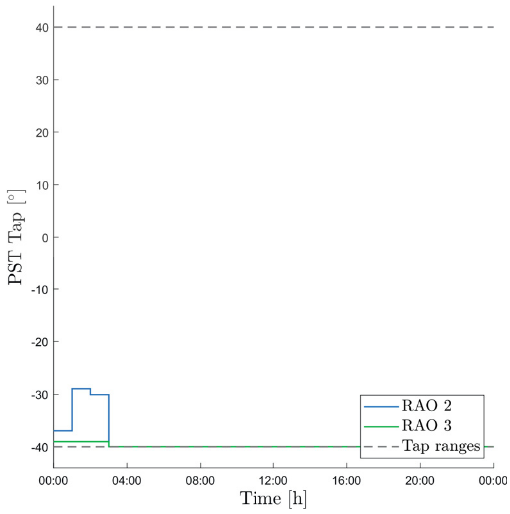
(a) PST in Rondissone (IT)



(b) PST in Camporosso (IT)



(c) PST in Foggia (IT)



(d) PSTs in Padriciano (IT), Divača (SI)

Figure 4.9: Optimized PST taps part 2

4.4.4 Conclusion

In this subsection, we summarize the previous results and compare the three different RAO model configurations that we investigate in this section. In total we compare the model configurations looking at reduced overloads, the amount of redispatch, and costs and at the PST tap settings.

When considering the overloads, RAO 2 and 3 are similarly good and outperform RAO 1. RAO 2 and 3 are also better than RAO 1 when investigating the total amount of redispatch and costs. Here, RAO 2 is better than RAO 3 and the best option. The cost advantage of configuration 2 above 3 has to be weighted against the results we see when

we investigate the PST tap settings. Overall, RAO 3 results in time series with less oscillations and tap settings that remain constant over longer periods of time compared to RAO 2. The savings of costs may not be of major advantage if they result in higher wear and tear of PSTs and costly renovations. To give a more sophisticated advice we need more data on the operation of PSTs and on their maintenance.
All in all, we can say that the coordinated usage of non-costly remedial actions like PSTs and HVDCs is of big advantage in remedial action optimization as it significantly reduces overloads, redispatch and costs.

5 Discussion

Remedial actions are important measures to attain operational security in the day-ahead coordination of transmission systems. Due to the way energy is traded on day-ahead markets it is possible that the ordered generations of power plants result in overloads on transmission lines. With information on the ordered net generation one can predict possible overloads and adjust accordingly. In this thesis we investigate different types of remedial actions and analyze their influence on the remedial action optimization. Doing so, our particularly interest lies in the benefits of the use of phase-shifting transformers and HVDC transmission which are non-costly remedial actions over the use of generator redispatch which is a costly remedial action.

To apply mathematical optimization algorithms to solve the problem of finding suitable remedial actions, we introduce in this thesis the concept of mathematical optimization and linear programming in particular. We describe a general linear optimization problem and investigate its properties. After we get a deeper insight into the problem structure, we introduce the class of interior-point methods which forms a family of algorithms to solve linear optimization problems. For these problems we state results that guarantee under mild assumptions the convergence of the algorithm to an optimal solution of the linear problem. Using these algorithms, we are able to tackle general linear optimization problems.

To use an interior-point method to solve the problem of remedial action optimization, we need to describe the problem using mathematical formalisms. To this end we introduce transmission systems in general and the physical equations that describe the power flow through a transmission system. We derive for these equations the DC load flow calculation which is a commonly used linearization that forms the backbone of our model. We further describe how we model the remedial actions that we use.

Before we model the problem of remedial action optimization, we describe the model scope and the data we use. We model the remedial action optimization for the European transmission system with focus on the region Italy North which contains the countries Austria, France, Italy, Slovenia, and Switzerland. The model covers a full day with one timestamp for each hour. We model power flows for the European grid but only allow remedial actions to be activated in the region Italy North. We only use open-source data mainly derived from the open-source software package PyPSA. We further define critical network elements that may fail during grid operation.

Therefore, we introduce the optimization algorithm we use, the equations to model the physical power flows and further the data sources and model scope. These are all ingredients we need to build a model for the remedial action optimization. The objective function we minimize contains the total costs that arise due to generator redispatch and overloads on transmission line elements. The model also takes into account line outages of the critical line elements we define beforehand.

To get a better insight in the remedial action optimization, we compare different model configurations. In the first model configuration we only allow the use of costly remedial actions. In the second model configuration we allow the use of all remedial actions. Both model configurations successfully decrease the total overloads on line elements. The use of non-costly remedial actions cuts the total overloads and costs in half compared to the sole use of generator dispatch. The analysis of the phase-shifting transformers reveals that their operation is impractical and the optimized tap positions of the phase-shifting transformers change from hour to hour. This can result in high wear and tear of the transformers. To remedy this problem, we develop a third model configuration with additional penalty terms in the objective function. This third model configuration performs with respect to costs and overloads as good as the second but we can see less oscillations in phase-shifting transformer tap settings.

All in all, in this thesis we successfully develop a model for remedial action optimization for the European transmission that is able to resolve overloads on transmission lines. We show the benefits of non-costly remedial actions over the pure use of generator redispatch. Using non-costly remedial action, we are able to significantly reduce overloads and costs. Additonally, we present a model configuration that can be used to regularize the operation of phase-shifting transformers.

List of Figures

List of Tables

References

[1] AIT GmbH. AIT Homepage. https://www.ait.ac.at/en/about-the-ait, 2021. Accessed: 2021-01-21.

[2] APG GmbH. APG Homepage. https://www.apg.at/Ueber-uns, 2021. Accessed: 2021-02-07.

[3] Council of the European Union European Parliament. Regulation (EC) no 1228/2003 of the european parliament and of the council of 26 june 2003 on conditions for access to the network for cross-border exchanges in electricity (text with EEA relevance). *Official Journal of the European Union*, 2003.

[4] Council of the European Union European Parliament. Regulation (EC) no 714/2009 of the european parliament and of the council of 13 july 2009 on conditions for access to the network for cross-border exchanges in electricity and repealing regulation (EC) no 1228/2003 (text with EEA relevance). *Official Journal of the European Union*, 2009.

[5] Council of the European Union European Parliament. Commission regulation (EU) 2015/1222 of 24 july 2015 establishing a guideline on capacity allocation and congestion management (text with EEA relevance). *Official Journal of the European Union*, 2015.

[6] Council of the European Union European Parliament. Regulation (EU) 2019/943 of the european parliament and of the council of 5 june 2019 on the internal market for electricity (text with EEA relevance). *Official Journal of the European Union*, 2019.

[7] Epexspot. Basics of the Power Market. https://www.epexspot.com/en/basicspower market#day-ahead-and-intraday-the-backbone-of-the-european-spot-market, 2021. Accessed: 2021-07-14.

[8] Council of the European Union European Parliament. Commission regulation (EU) 2017/1485 of 2 august 2017 establishing a guideline on electricity transmission system operation (text with EEA relevance). *Official Journal of the European Union*, 2017.

[9] European Union Agency for the Cooperation of Energy Regulators. Methodology for regional operational security coordination for the core ccr in accordance with article 76(1) of commission regulation (EU) 2017/1485 of 2 august 2017 establishing a guideline on electricity transmission system operation: Annex I. 2020.

[10] Dantzig, G. B. *Origins of the Simplex Method*. Association for Computing Machinery, 1990.

[11] Geiger, C. and Kanzow, C. *Theorie und Numerik Restringierter Optimierungsaufgaben*. Springer Berlin Heidelberg, 2002.

[12] Wright, S. J. *Primal-dual interior-point methods.* SIAM, 1997.

[13] Klee, V. and Minty, G. J. How good is the simplex algorithm. *Inequalities*, 3(3):159–175, 1972.

[14] Schwab, A. J. *Elektroenergiesysteme, Verbundsysteme.* Springer, 2020.

[15] Wood, A. J., Wollenberg, B. F., and Sheblé, G. B. *Power generation, operation, and control.* John Wiley & Sons, 2013.

[16] Verboomen, J. *Optimisation of transmission systems by use of phase shifting transformers.* PhD Thesis, Delft University of Technology, 2008.

[17] Van Hertem, D., Verboomen, J., Cole, S., Kling, W., and Belmans, R. Influence of phase shifting transformers and hvdc on power system losses. *Proceedings of PES 2007*, 2007.

[18] Van Hertem, D., Verboomen, J., Purchala, K., Belmans, R., and Kling, W.L. Usefulness of dc power flow for active power flow analysis with flow controlling devices. In *The 8th IEE International Conference on AC and DC Power Transmission*, pages 58–62, 2006.

[19] Powell, L. *Power System Load Flow Analysis.* McGraw-Hill Education, 2004.

[20] Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. Julia: A fresh approach to numerical computing. *SIAM review*, 59(1), 2017.

[21] PyPSA-Eur: An Open Optimisation Model of the European Transmission System. https://pypsa-eur.readthedocs.io/en/latest/, 2021. Accessed: 2021-04-23.

[22] Brown, T., Hörsch, J., and Schlachtberger, D. PyPSA: Python for Power System Analysis. *Journal of Open Research Software*, 6(4), 2018.

[23] Van Rossum, G. and Drake, F. L. *Python 3 Reference Manual.* CreateSpace, 2009.

[24] IBM ILOG CPLEX. V12. 1: User's manual for CPLEX. *International Business Machines Corporation*, 46(53), 2009.

[25] J. Hörsch, F. Hofmann, D. Schlachtberger, and T. Brown. Pypsa-eur: An open optimisation model of the european transmission system. *Energy Strategy Reviews*, 22:207–215, 2018.

[26] PyPSA-Eur: An Open Optimisation Model of the European Transmission System (Dataset). https://zenodo.org/record/3601882#.YOglBOgzaUl. Accessed: 2021-04-28.

[27] Novák, S. *Využití phase-shift transformátorů v přenosových soustavách.* Bachelor's Thesis, Czech Technical University in Prague, 2017.

[28] Tesař, J. *Koordinace provozu PST transformátorů v propojeném regionu.* Master's Thesis, Brno University of Technology, 2018.

[29] Bresesti, P., Sforna, M., Allegranza, V., Canever, D., and Vailati, R. Application of phase shifting transformers for a secure and efficient operation of the interconnection corridors. *IEEE Power Eng Soc General Meeting*, 2:1192 – 1197 Vol.2, 07 2004.

[30] FARAO community. FARAO: Fully Autonomous Remedial Actions Optimisation (Documentation). https://farao-community.github.io/docs/, 2021. Accessed: 2020-01-12.