

École polytechnique de Louvain

Economic Emission Dispatch

Author: **Félix MOREL DE WESTGAYER**

Supervisor: **Pierre-Antoine ABSIL**

Readers: **Anthony PAPAVALIOU, Loïc VAN HOOREBEECK**

Academic year 2019-2020

MAP2M: Master [120] in Mathematical Engineering

Abstract

The Economic Emission Dispatch (EED) is an important multi-objective optimization problem in fossil fueled electric power plants operation. It consists in allocating the power production of each committed generator while respecting the operational constraints. The objective of this problem is to minimize simultaneously the total fuel costs and the pollutant emissions that are related to the combustion of the fuels. The EED is a highly nonlinear constrained optimization problem and this thesis focuses on deterministic optimization algorithms to solve the problem. In order to achieve this, different features of the problem are studied. Of all models studied, the most complex one treated in this thesis consists of the static EED using a nonlinear emission function considering the transmission losses and the Prohibited Operating Zones (POZ). Examples from the literature are included in this work in order to validate the results and to analyze in detail the impact of the emissions on the problem.

Acknowledgments

This master's thesis represents the accomplishment of my studies at the *École Polytechnique de Louvain* graduating as a master in Mathematical Engineering. This work was supported by professor Pierre-Antoine Absil and PhD student Loïc Van Hoorebeeck, who is currently working on the Economic Dispatch Problem.

Firstly, I would like to thank my thesis supervisor, Pierre-Antoine Absil, for his frequent reviews throughout this semester and his genuine advice for completing my work. In spite of the total lockdown due to the Covid-19 pandemic during the major part of my work, he really stayed available and easy to reach.

Then, I would particularly like to express my gratitude to Loïc Van Hoorebeeck for his strong investment in my thesis. Being really available and engaged, he often offered his help for addressing questions, revisions and critics about my work. His contributions have been essential to achieve the best outcome of my research.

Finally, I would like to thank my family for the continuous moral and logistical support. Their presence has been invaluable to me and I would like to thank them for their ever-lasting trust they had in me.

Contents

Introduction	1
1 Pareto optimal front for a simple static problem	3
1.1 Model	3
1.2 Pareto optimal front	4
1.2.1 Kalai-Smorodinsky solution	5
1.3 Scalarization Method	6
1.3.1 Analysis of the weight on the scalarized objective	7
1.3.2 Smart spacing of the weights	8
1.4 ϵ -constraint method on the Emission	10
1.5 ϵ -constraint method on the Costs	11
1.6 Piecewise Hermite approximation	13
1.6.1 Error analysis	14
2 Dynamic problem	16
2.1 Dynamic demand and ramp rates	16
2.2 Reserve Constraints	18
3 Nonlinear Emission objective	20
3.1 Exponential term in the emissions	20
4 Transmission losses	22
4.1 Non-convex problem	22
4.2 Pareto front analysis	23
4.2.1 Scalarization method	23
4.2.2 ϵ -constraint method	24
4.3 Gradient based methods on non-convex programming	26
4.3.1 Constrained gradient descent	26
4.3.2 Accelerated Gradient	27
4.4 Analysis of the achievable domain of objectives	28
4.4.1 2 generator analysis	28
4.4.2 3 generator analysis	30
4.4.3 Generalisation	31
4.5 Approach to relax the non-convex demand	31
4.5.1 (Hyper)plane constructed with N extreme points	32
4.5.2 Feasible set: convex relaxation	33
4.5.3 Feasible set: Linear relaxation	35
4.6 Relaxation problems	36
4.6.1 Convex Relaxation	36

4.6.2	Linear Relaxation using N extreme points	41
4.6.3	Linear relaxation from 1 feasible point	45
4.7	Comparison of the relaxations	46
4.7.1	Error analysis	46
4.7.2	Computational time	49
4.8	First order solvers	51
4.8.1	Gradient and Accelerated Method	51
4.8.2	Spectral Projected Gradient	52
4.9	Second order methods	54
4.9.1	Sequential Quadratic Programming	55
5	Prohibited Operating Zones	57
5.1	Problem formulation	57
5.2	Scalarization of the objective	58
5.3	ϵ -constraint method	61
	Conclusion	64
A	Test cases	66
A.1	2-unit case	66
A.2	3-unit case	66
A.3	6-unit case	66
A.4	10-unit case	66
A.5	40-unit case	66
A.6	100-unit case	67

Notations

Variables:

N	Number of generating power plants.	
$\mathbf{p}$	Set of outputs, vector of size N .	
p_i	Output of the i^{th} unit.	
$\mathbf{P}^{\min}$	Vector of the lower limits for all units.	
$\mathbf{P}^{\max}$	Vector of the upper limits for all units.	
$F(\mathbf{p})$	Fuel cost function.	[\$]
$E(\mathbf{p})$	Emission function.	[lb]
a_i, b_i, c_i	Fuel cost coefficients of the i^{th} unit.	
$\alpha_i, \beta_i, \gamma_i, \delta_i, \eta_i$	Emission coefficients of the i^{th} unit.	
P_D	Load demand.	[MW]
B	Transmission Loss matrix.	

Sets:

$\mathbb{R}^N$	N -dimensional real space.
$\mathbb{R}_+^N$	N -dimensional nonnegative real space.
$[\mathbf{P}^{\min}, \mathbf{P}^{\max}]^N$	Feasible box domain defined by $[P_1^{\min}, P_1^{\max}] \times [P_2^{\min}, P_2^{\max}] \times \dots \times [P_N^{\min}, P_N^{\max}]$.

Operator:

$\mathbf{p} \preceq \mathbf{q}$	Vector $\mathbf{p}$ is dominated by $\mathbf{q}$: $p_i \leq q_i$	$i = 1, \dots, N$
---------------------------------	---	-------------------

Introduction

Climate change is already affecting our entire world [1]. The global warming leading to extreme weather conditions has drastic consequences on ecosystems, economic sectors and people's health and well-being. The climate change is a worldwide challenge. As examples, the EU's objective is to reduce its greenhouse gas emission by 40% in 2030 compared to 1990, and the Paris Agreement, signed by 195 countries, provides the target to attain the carbon neutrality by mid 21st century [2]. These ambitious climate policies require drastic changes of our society.

According to the Global Carbon Project, it has been reported that the global annual emissions of CO_2 stabilizes slowly but continues to increase by 0.6%, summing up to 37 Gigatons of CO_2 in 2019 [3]. On the other hand, natural carbon sinks - systems that are able to absorb more carbon than they produce, like forests, soil and oceans - remove between 9.5 and 11 Gt of CO_2 per year according to Le Quéré et al. [4]. Currently, there is a big gap between the total emissions and the absorbed carbon. It is therefore essential to reduce carbon emission in order to reach carbon neutrality. The Global Carbon Project evaluated that 45% of the total CO_2 emissions due to the combustion of carbon are from the energy sector. In fact, since 61% of worldwide's electricity is produced by fossil fueled power plants, they are the main contributors of this global warming [5]. Indeed, NO_x , SO_2 and CO_2 are the major pollutants from thermal power plants due to the combustion of fossil fuel [6].

The Economic Dispatch (ED) is a well studied problem in the scientific world, consisting of an optimization problem for allocating the power generation by fossil fueled power plants while minimizing its generation costs. At a more detailed level, the Economic Emission Dispatch (EED) is a branch in the ED field and takes the environmental impact into account within the ED to decide the allocation between each power plant. With the increasing and indispensable concern of pollution, thermal power plants will have to reduce their emissions generated by the combustion of fossil fuel.

This thesis studies one of the alternative reductions solutions, which do not require any capital investments on existing power plants or any modifications in the process. This short-term alternative consists of dispatching the required power generation while minimizing emissions and cost. The goal of this problem is therefore to optimize simultaneously multiple objectives and is known as a multi-objective optimization program.

The ED problem is already widely discussed in the literature [7], being characterized as a large-scale highly nonlinear constrained optimization program. Currently, a large number of heuristics are applied in order to solve the problem in a relatively short time period. In addition, many papers approach the EED by using heuristic methods for multi-objective optimization (see, e.g., [8], [9], [10] and [11]). Heuristics are very attractive methods due to their time efficiency while giving reasonable solutions, but do not provide any guarantee regarding the quality level

of their returned solutions.

On the other hand, this thesis will mainly focus on the application of deterministic algorithms for solving this problem. Even though this area is less developed in the literature, some papers are approaching the EED problem on this side as explained in [7]. Deterministic mathematical optimization programs present their strength in the quality of the solutions but are poor in computational time when the nonlinearity and discontinuous characteristics are considered. Therefore, first order solvers, namely the Constrained Gradient Method, the Accelerated Gradient and the Spectral Projected Gradient (SPG) are applied on the EED in this thesis in order to verify their efficiency. In addition, the Sequential Quadratic Programming (SQP), a second order solver, is also analyzed in this study.

In addition, the other main objective of this master's thesis is to analyze the impact of the emissions on the EED problem. It has been observed that the minimal cost solution does not provide the minimum possible emission. The complexity of multi-objective optimization is to find a trade-off between these two conflicting objectives.

This master's thesis is structured by increasing complexity of the model. The first chapter describes the simplest mathematical formulation of the model that consists in the static version defined by a quadratic problem (QP) with linear constraints. In addition, basic notions of multi-objective optimization are explained and illustrated by means of this model.

The second chapter tackles the dynamic model of the EED by taking the ramp rate limits of the generators into account. The dynamic problem is introduced and considers a fluctuating demand in function of time.

Chapter 3 presents a better expression of the emission function resulting in a nonlinear convex optimization.

The fourth chapter consists in considering the losses in the transmission lines. This leads to a non-convex optimization program. Relaxations and algorithms are studied in order to efficiently solve the EED problem with the transmission losses.

Finally, the last chapter describes and analyzes the problem that additionally handles the Prohibited Operating Zones (POZ) of the power plants. This additional feature leads the problem into a Mixed Integer Nonlinear Program (MINLP) that has to be solved.

The main contributions of this master's thesis are therefore mainly developed in chapter 4 and 5 of this report. Indeed, until now, scientific papers tackled the EED with heuristics, namely evolutionary algorithms, or hybrid methods that combines several heuristic techniques and deterministic approaches [9]. In this thesis, optimization algorithms are used and analyzed in order to measure their performances. In addition, new relaxations on the EED are proposed and analyzed in this thesis.

Finally, the set of solutions associated to EED are obtained and analyzed in this report given test systems inspired from the literature. Therefore, a specific detailed analysis of the Pareto optimal front is realized for each of the considered problems in order to evaluate the impact of emissions in the problem.

Chapter 1

Pareto optimal front for a simple static problem

This first chapter has for objective to describe a simple model for approaching the Economic Emission Dispatch. The main goal is to present an introductional multi-objective optimization program and to describe the basic notions for solving these problems. The major reference for this chapter is taken from *Convex optimization* by Stephen Boyd and Lieven Vandenberghe (see [12]).

The first section of this chapter describes the model in its most essential form. Then, the Pareto optimal front (POF) and other concepts of multi-objective optimization that will be used further in this report are defined in the following section. In order to compute this POF, the next part has for aim to explain and analyze the scalarization method on several examples. In addition, the ϵ -constraint method generates another form of optimization program in order to browse the Pareto optimal values. Section 1.4 and 1.5 describes this method by limiting the emission rate and the fuel costs respectively. Finally, the last section specifies the Hermite interpolation in order to construct a continuous approximation of the Pareto front.

1.1 Model

Firstly, this problem has to be modelled into an optimization program. As a first step, the static problem is considered. This simplified formulation makes the definition of the concepts introduced in this chapter easier. The multi-periodic problem, also known as dynamic problem, is tackled in chapter 2.

The static problem consists in having a single power demand P_D , expressed in MW, at one period of time. The problem is defined as finding the power production p_i for each unit while minimizing both fuel costs and the rejection of pollutants. We call $\mathbf{p}$ the stacked vector of all productions. Our problem consists in N generators, each having a particular cost function and emission function. They can be modeled in the first time as quadratic functions. A more complex approximation of the emission function is studied in chapter 3. The other constraints consist in the generation capacity of each unit. The initial model has therefore the following form [9].

$$\begin{aligned}
 \min_{\mathbf{p}} f : & \begin{cases} F(\mathbf{p}) = \sum_{i=1}^N (a_i + b_i p_i + c_i p_i^2) & [\text{Fuel costs}] \\ E(\mathbf{p}) = \sum_{i=1}^N (\alpha_i + \beta_i p_i + \gamma_i p_i^2) & [\text{Emissions}] \end{cases} \\
 \text{s.t. } \mathbf{p} \in \mathcal{D} : & \begin{cases} \sum_{i=1}^N p_i = P_D & [\text{Demand}] \\ P_i^{\min} \leq p_i \leq P_i^{\max} & i = 1, \dots, N \quad [\text{Limits}] \end{cases}
 \end{aligned} \tag{1.1.1}$$

The main assumptions are that parameters $a, b, c, \alpha, \gamma \in \mathbb{R}_+^N$, $\beta \in \mathbb{R}^N$ and that $\sum P_i^{\min} \leq P_D \leq \sum P_i^{\max}$ in order to be feasible. This problem is known as a multi-objective optimization problem. The difficulty of having multiple objectives is that each objective function has its own minimum and usually these optimums are not located at the same solution. In addition, the complexity is that the two objectives work with different units. The fuel costs are measured in US dollars while the emissions represents the total mass of all pollutants given in pounds. We therefore must find a trade-off to combine properly these two objectives.

1.2 Pareto optimal front

Unlike for single objective optimization problems, a multi-objective optimization program has in general no unique optimal solution anymore. The notion of Pareto optimality can be introduced. The problem can be written as a vector optimization problem of the form:

$$\begin{aligned}
 \min \mathbf{f}_0(\mathbf{x}) \\
 \text{s.t. } f_i(\mathbf{x}) \leq 0 \quad i = 1, \dots, 2N \\
 h(\mathbf{x}) = 0
 \end{aligned} \tag{1.2.1}$$

where $\mathbf{f}_0(\mathbf{x}) = \begin{bmatrix} E(\mathbf{p}) \\ F(\mathbf{p}) \end{bmatrix}$, $f_i(\mathbf{x})$ corresponds to the limit constraints and h corresponds to the demand.

Given a general multi-objective optimization problem of n objectives, m inequality constraints and p equality constraints, the set of achievable objective values can be defined as:

$$\mathcal{O} = \left\{ \mathbf{f}_0(\mathbf{x}) \mid \exists \mathbf{x} \in \mathbb{R}^N, \begin{matrix} f_i(\mathbf{x}) \leq 0 & i = 1, \dots, m \\ h_i(\mathbf{x}) = 0 & i = 1, \dots, p \end{matrix} \right\} \subseteq \mathbb{R}^n$$

This problem does not necessarily have a minimum value x such that $\mathbf{f}_0(\mathbf{x}) \preceq \mathbf{f}_0(\mathbf{y})$ for all feasible $\mathbf{y}$. The operator $\succeq$ states that every element of the vector has to respect the inequality. It is defined that $\mathbf{p}$ dominates $\mathbf{q}$ if $\mathbf{p} \succeq \mathbf{q}$. The minimal elements of $\mathcal{O}$ that are called Pareto optimal values can therefore be analyzed. The definition states [12]: *A point x is Pareto optimal if and only if it is feasible and*

$$(\mathbf{f}_0(\mathbf{x}) - K) \cap \mathcal{O} = \{\mathbf{f}_0(\mathbf{x})\}$$

with $K \in \mathbb{R}_+^n$ and the set $\mathbf{f}_0(\mathbf{x}) - K$ can be interpreted as the set of values that are better or equal to $\mathbf{f}_0(\mathbf{x})$. The Pareto optimal front (POF) consists of the set of Pareto optimal values. These values lie on the boundary of the set of achievable objective values as illustrated on the

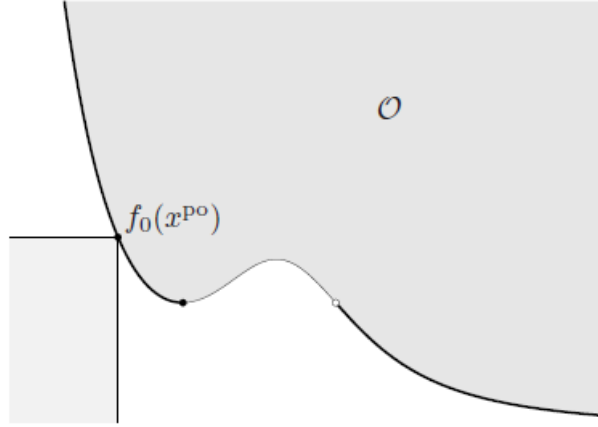


Figure 1.2.1: Set $\mathcal{O}$ of achievable values for a vector optimization problem with objective values in $\mathbb{R}^2$.

darkened curve in Figure 1.2.1. The Pareto optimal point x^{po} can be observed with the shaded box on the right that corresponds to $\mathbf{f}_0(x^{po}) - K$.

For bi-objective optimization problems like the Economic Emission Dispatch problem, it is common to display the objective values on a graph that is called the feasible objective space. Each axis represents one objective value and the graph can be perceived as a parametric region obtained by the solutions $\mathbf{x}$. The POF is therefore plotted on this graph to get a geometric visualization.

1.2.1 Kalai-Smorodinsky solution

Since multi-objective optimization problems generally do not provide a unique solution due to the conflicting objectives, Kalai-Smorodinsky (KS) proposed a well-balanced solution for multi-objective problems with no preference on any objective function. This point can therefore be interesting to use in the case a solution has to be chosen and gives a good compromise between the objectives.

From [13], this solution is obtained from the Pareto front and its extreme values. Let us define the point of best utilities U_t that corresponds to the minimal achievable value of the objectives. On the other side, we can determine the point of disagreement D that is equal to the worst value of each objective on the POF. The KS solution consists in the intersection point between the segment created by U_t and D with the frontier of the achievable set of objectives as illustrated on figure 1.2.2 for a bi-objective problem.

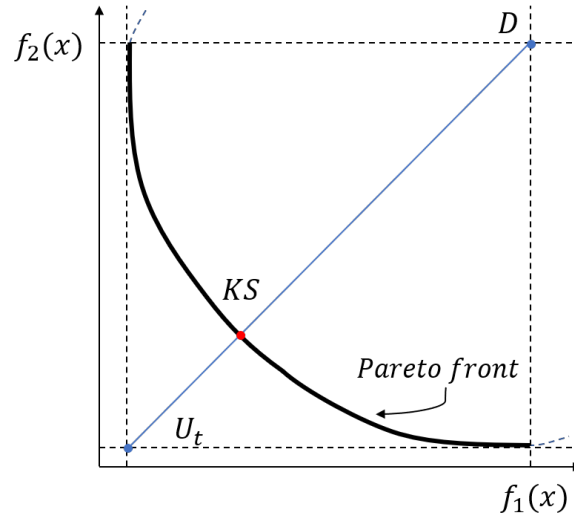


Figure 1.2.2: Solution of Kalai-Smorodinsky (KS) between two functions f_1 and f_2 .

1.3 Scalarization Method

The weighted scalarization of the objective function is a technique for obtaining Pareto optimal points. It consists in using a weighting vector $\lambda \in \mathbb{R}_{++}^n$ where $\mathbb{R}_{++}^n$ is defined by the set of vectors of size n such that every element is positive. This weighting vector converts the vector optimization problem in a scalar optimization problem:

$$\begin{aligned} \min \quad & \lambda^T f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0 \quad i = 1, \dots, m \\ & h_i(x) = 0 \quad i = 1, \dots, p. \end{aligned} \tag{1.3.1}$$

Moreover, if the set is nonempty closed convex and $\lambda^T f_0$ is a strictly convex function on its domain for some nonzero $\lambda \succeq 0$, the solution x is unique [12]. This statement allows therefore to work with $\lambda \succeq 0$ such that $\lambda \neq 0$. In addition, if the problem is strictly convex, then every Pareto solution is attainable by the scalarization method.

In order to find a range of Pareto optimal points in this specific problem, we will firstly work with the following scalarized objective function:

$$\min_p \quad \lambda^T f_0(p) = wE(p) + (1-w)F(p) \quad w \in [0, 1] \tag{1.3.2}$$

Let us therefore apply the scalarization method on some test cases in order to obtain their POF. The three main test cases that will be used in this report are given in Appendix A.3, A.4 and A.5 and are called 6-unit test case, 10-unit test case and 40-unit test case. Throughout this document, the model will be complexified and its associated parameters will extend this data. However, we should stay critical on this data because some inconsistencies are present. Indeed, according to the literature, the 40-unit test case has some generators that produces a negative emission rejection. In order to compare the figures with the literature, it has been decided to keep this data unchanged. Note that even if it has no physical sense, it does not troubles the optimization program.

On Figure 1.3.1 is displayed in red the Pareto optimal values obtained by varying the weight w from 0 to 1. The first problem on the left on the figure corresponds to a test case with 6

generators. The second deals with 10 generators and the last one has 40 generators. As it can be noticed, the POF is displayed on a graph with the emission load given in abscissa and the fuel costs as ordinate. The figures show that the resulting front is convex and that the optimal values are not spaced equally.

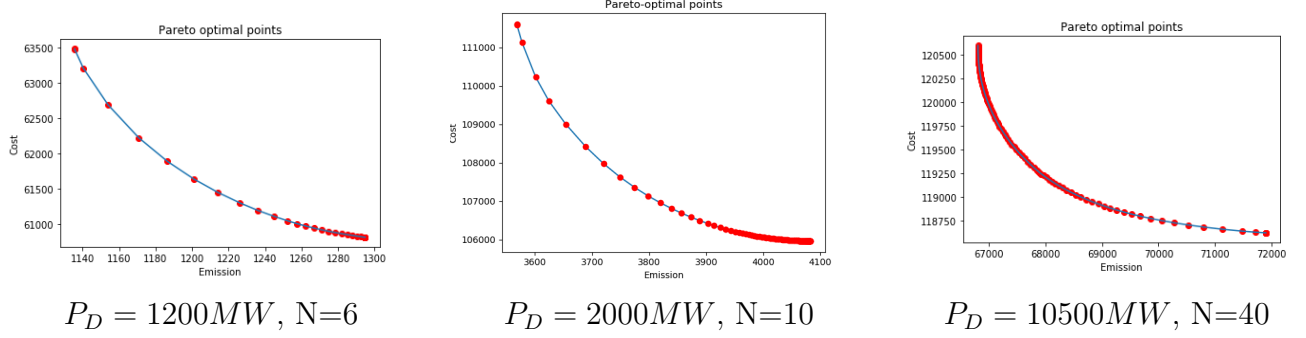


Figure 1.3.1: Pareto optimal front for problems of different demands and number of units.

1.3.1 Analysis of the weight on the scalarized objective

From the latter section, it has been shown that each different weight λ corresponds to a unique Pareto optimal point. It is therefore interesting to study the influence of λ on the location of the solution on the POF. This subsection explains that there exists a geometrical relationship between vector λ and the Pareto front: namely that the vector is tangent at this curve.

We know that for each weight, we find a unique optimal solution x that minimizes $\lambda^T f_0(x)$. In other words, for all feasible $y \neq x$, $\lambda^T(f_0(y) - f_0(x)) > 0$. In this problem, it leads to:

$$\begin{aligned}
 &(\lambda_1, \lambda_2) \left(\begin{bmatrix} E(y) \\ F(y) \end{bmatrix} - \begin{bmatrix} E(x) \\ F(x) \end{bmatrix} \right) > 0 \\
 \Leftrightarrow &\lambda_1[E(y) - E(x)] + \lambda_2[F(y) - F(x)] > 0 \\
 \Leftrightarrow &F(y) > -\frac{\lambda_1}{\lambda_2}[E(y) - E(x)] + F(x)
 \end{aligned}$$

This expression can be seen as the equation of the tangent at $E(x)$ in the feasible objective space with slope equal to $-\frac{\lambda_1}{\lambda_2}$. Geometrically, we can represent the weight vector as a search direction in the $(E(p), F(p))$ graph. By taking the set of perpendicular planes to λ , the intersection between the POF and its tangent having a direction of $\lambda^\perp$, results to the Pareto optimal values of the problem. We can therefore conclude that if the problem is convex, then the POF will always be convex. Figure 1.3.2 illustrates the match between the weight vector (green) and the Pareto curve tangent (blue).

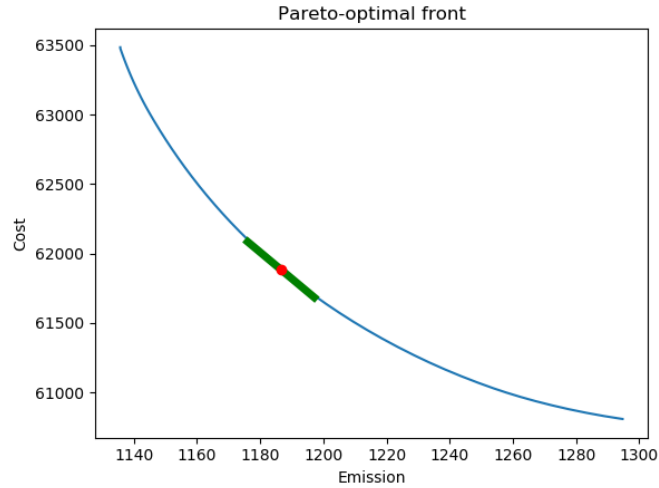


Figure 1.3.2: Tangent of search direction with $N=6$.

1.3.2 Smart spacing of the weights

From the graphs of Figure 1.3.1, it is noticed that the objective values from the Pareto optimal points are often grouped at one extremity on the curve. The reason is that the two objectives do not work at the same order of magnitude leading to group these values close to the optimum of the objective with large variations. It should therefore be interesting to focus on a technique to spread equally the computed optimums on the front in order to limit the number of computations and to increase the relevance of each point.

In order to achieve this, we will work with a scalarized objective that has the form:

$$\min_{\mathbf{p}} \pi w E(\mathbf{p}) + (1 - w) F(\mathbf{p}) \quad w \in [0, 1]. \quad (1.3.3)$$

Multiplier π represents an average price on the emissions that producers have to pay in order to have an objective function that represent all of the costs. In such a way, we compare effectively the cost of production and the costs on the emissions.

Modified Price Penalty Factor

The literature proposes a factor that helps to balance the two objectives [10]. This factor, noted by h_m , consists in an estimation of the price on the emissions. This can be computed without solving any optimization program in advance. The algorithm for obtaining this factor is given by:

Algorithm 1 Modified Price Penalty Factor

- 1: Given load demand P_D : $\sum P_i^{min} \leq P_D \leq \sum P_i^{max}$.
 - 2: $h_i := \frac{F(P_i^{max})}{E(P_i^{max})}$ for each i .
 - 3: $index := argsort(\mathbf{h})$
 - 4: $\mathbf{h} \leftarrow \mathbf{h}[index]$
 - 5: $\mathbf{m} := \mathbf{P}^{max}[index]$
 - 6: $P=0, k=1$
 - 7: **while** $P + m_k < P_D$ **do**
 - 8: $P \leftarrow P + m_k$
 - 9: $k \leftarrow k + 1$
 - 10: **end while**
 - 11: $\nu = \frac{P_D - P}{m_k}$
 - 12: $h_m = (1 - \nu)h_{k-1} + \nu h_k$
-

The scalarized objective can be formulated with parameter h_m as:

$$\min_{\mathbf{p}} h_m w E(\mathbf{p}) + (1 - w) F(\mathbf{p}) \quad w \in [0, 1]$$

When we compute the optimal values for the same range of weights as for Figure 1.3.1, we obtain now the results illustrated in Figure 1.3.3. The points are better spaced but it is not ideal yet. Obviously, the 40-unit test case does not have all its points perfectly equally spaced. This is because the price is overestimated since the algorithm computes a price for a feasible solution but that is not optimal.

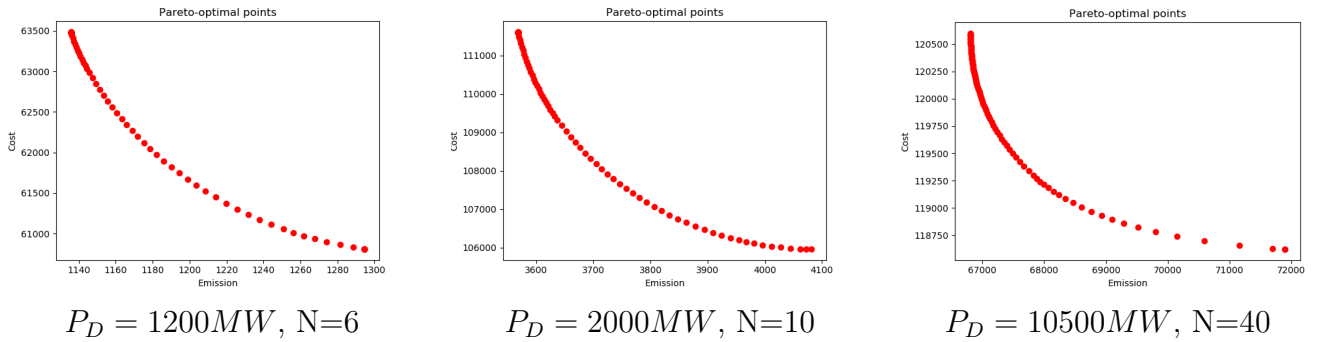


Figure 1.3.3: Pareto optimal front with better spacing using modified price penalty factor.

Average slope

Another way to decide on a price on the emissions can be by considering the solution of the economic and the environmental dispatch problem. The solution of the economic dispatch corresponds to setting $w = 0$ and the emission dispatch when $w = 1$. With these two optimums, we can compute the average slope of the POF and use this to fix an average price on the emissions.

Let us define the economic power production dispatch as $\mathbf{p}^{Eco}$ and $\mathbf{p}^{Env}$ for the environmental problem. The average price is given by:

$$\pi = \frac{F(\mathbf{p}^{Env}) - F(\mathbf{p}^{Eco})}{E(\mathbf{p}^{Eco}) - E(\mathbf{p}^{Env})}. \quad (1.3.4)$$

The results of this method are much more convincing than the latter as we can see on Figure 1.3.4. It allows therefore to compute less Pareto optimal points but unlike the modified price penalty factor, this technique requires to solve 2 optimization problems to obtain the average price. However, since these two optimization problems give valid Pareto optimal values, they would be computed in either way. We will therefore work with this improved method when computing the Pareto optimal values using the scalarization of the objective.

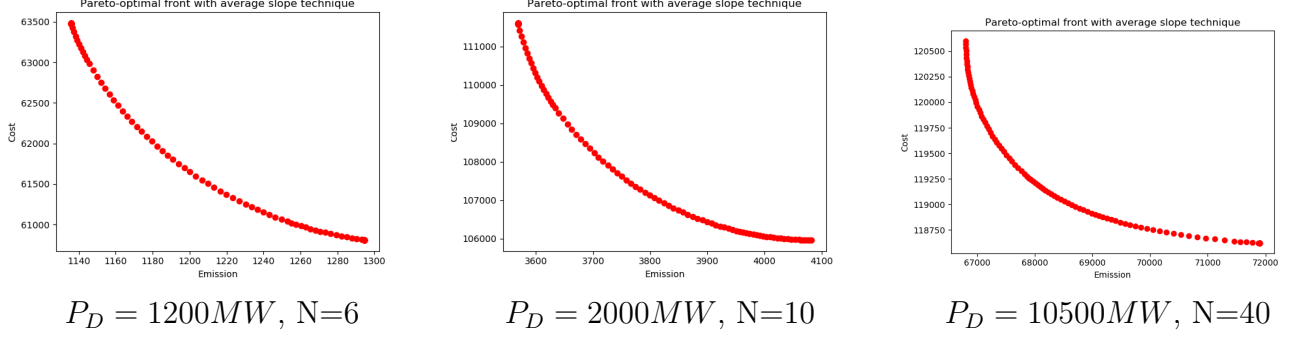


Figure 1.3.4: Pareto optimal front with better spacing using the average slope.

1.4 ϵ -constraint method on the Emission

Let us consider that the government wants to impose a limitation on the maximum emission rejection. They could indeed impose a restriction on the maximum emission or inversely, fix a tax on the pollution. By knowing in advance the parameters of the problem, it should indeed be possible to determine a tax on the emission in order to restrict implicitly the maximum emissions generated by the power plants. By imposing a penalty on the emissions, the fossil fueled power plants will have to produce less emissions in order to minimize their total costs. This section works with an alternative of the scalarization method for obtaining the POF that is the ϵ -constraint method. The resulting applications of this method are also explained.

The ϵ -constraint method for a multi-objective optimization problem consists in minimizing one objective function while constraining the others. For the EED problem, it is therefore possible to minimize the fuel costs and constrain the emissions, or inversely. Let us first use the ϵ -constraint method on the emissions. The optimization model can therefore be written as:

$$\begin{aligned}
 \min_{\mathbf{p}} F(\mathbf{p}) &= \sum_{i=1}^N (a_i + b_i p_i + c_i p_i^2) \\
 \text{s.t. } E(\mathbf{p}) &= \sum_{i=1}^N (\alpha_i + \beta_i p_i + \gamma_i p_i^2) \leq E_{max} \quad (\lambda) \\
 \sum_{i=1}^N p_i &= P_D \\
 P_i^{min} &\leq p_i \leq P_i^{max} \quad i = 1 \dots N.
 \end{aligned} \tag{1.4.1}$$

The dual variable λ of the emission-constraint will be an important parameter later in the text. Graphically, the optimum will lie on the intersection between the POF and the limit on the emissions E_{max} . By varying parameter E_{max} , it is possible to get all the points on the POF.

In addition, the price on the emissions (in $[\$/lb]$) can be obtained, let us call it λ_E . This tax price on the emissions should therefore be established in order to force the power plants to generate less emissions than the fixed ceiling E_{max} for a given demand P_D . Indeed, given the price on the emissions, the objective function that the power operator would have to minimize is given by:

$$\min_{\mathbf{p}} F(\mathbf{p}) + \lambda_E E(\mathbf{p}). \quad (1.4.2)$$

This problem consists in the Lagrangian relaxation of problem (1.4.1) by relaxing the constraint on the emissions. Since the problem is convex and given E_{max} in the interval defined with $E(\mathbf{p}^{Env})$ and $E(\mathbf{p}^{Eco})$, the emission-constraint is always tight leading to an exact relaxation. On this interval, the optimum will lie on the POF and since the front is continuous, the objective values will lie at the intersection between the POF and the limit on emissions E_{max} . The dual variable of the ϵ -constraint is therefore equal to the opposite of the price or:

$$\lambda_E = -\lambda.$$

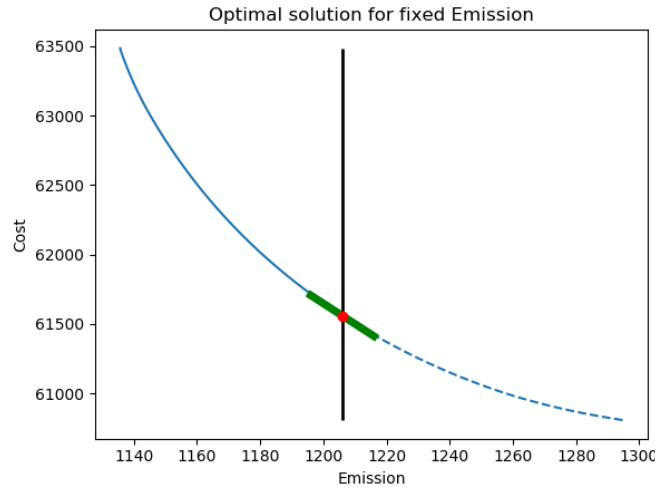


Figure 1.4.1: Fixed Emission at $E_{max} = 1206$ lb with $N=6$.

As it can be noticed from Figure 1.4.1, adding the ϵ -constraint on the emissions limits the attainable objective values on the left of the black line corresponding to the value of E_{max} . Since the optimization program minimizes the costs, the objective values associated to the minimum is illustrated by the red dot in the feasible objective space. In order to validate the coherence of the computed price, the tangent associated to this price is computed and displayed by the green line. As expected from section 1.3.1, it corresponds indeed to the gradient of the Pareto front.

As an example, by fixing a maximum emission of 1206 lb on the 6-unit test system, the program computes that the recommended price is $\lambda_E = 14.71$ $\$/lb$. From Figure 1.4.1, it seems indeed correct when we look at the green line that has a slope of -14.71.

1.5 ϵ -constraint method on the Costs

The same work can also be performed if the power distributor has a certain maximal cost limit that it cannot exceed. The cost is therefore fixed and the problem consists in reducing the

emissions. The optimization problem becomes:

$$\begin{aligned}
 \min_{\mathbf{p}} E(\mathbf{p}) &= \sum_{i=1}^N (\alpha_i + \beta_i p_i + \gamma_i p_i^2) \\
 \text{s.t. } F(\mathbf{p}) &= \sum_{i=1}^N (a_i + b_i p_i + c_i p_i^2) \leq C_{max} \quad (\lambda) \\
 \sum_{i=1}^N p_i &= P_D \\
 P_i^{min} &\leq p_i \leq P_i^{max} \quad i = 1 \dots N.
 \end{aligned} \tag{1.5.1}$$

As in the latter section, the dual variable λ of the cost-constraint has to be computed in order to obtain the equivalent problem with the weighted objective function given by:

$$\min_{\mathbf{p}} E(\mathbf{p}) + \lambda_{Fuel} F(\mathbf{p}). \tag{1.5.2}$$

Here $\lambda_{Fuel} = -\lambda$ for the same reasons as in the previous section. The gradient at the optimum can also be computed and is equal to $\frac{1}{\lambda}$. Graphically, as we can see in Figure 1.5.1, the attainable objective space is defined by the region below the limit on the cost, represented by the black line, and above the POF. Since the problem is to minimize the emissions, we find the unique optimum given by the red dot.

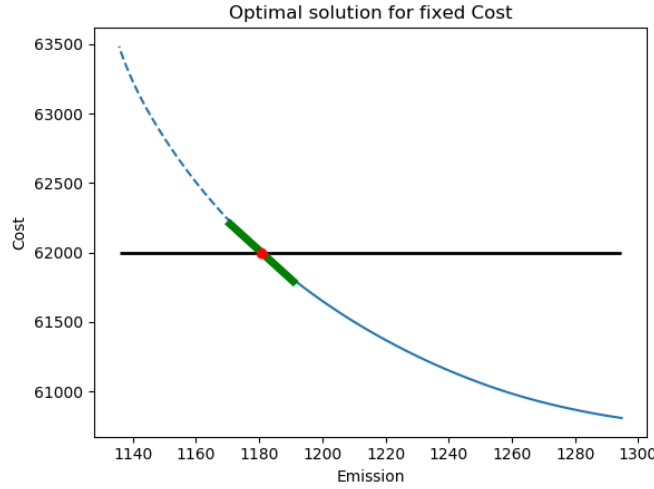


Figure 1.5.1: Fixed Cost of $C_{max} = 62k$ \$ with $N=6$.

As a conclusion, the ϵ -constraint method is able to recover the whole Pareto front. In contrary to the scalarization of the objective, this technique does not require to have a convex Pareto front in order to get all Pareto values. The drawback of this method is that it is required to know the range in which the Pareto optimal values lie in order to constrain them with adequate limits. In addition, the method is not accurate for obtaining the points on the front where its slope is close to be horizontal or vertical depending which objective is constrained. For example, small variations in the emission decreases hugely the fuel costs in the region near the Emission Dispatch objective values (on the top corner left) if we consider the ϵ -constraint method on the emissions. A good implementation based on the ϵ -constraint method should therefore be to constrain the fuel costs near the Emission Dispatch and inversely, constrain the emissions when approaching the Economic Dispatch optimum.

1.6 Piecewise Hermite approximation

The scalarization method or the ϵ -constraint method allow to obtain for each problem a single Pareto optimal value. Hence, only a discrete front can be obtained by solving a finite number of problems. In order to have a continuous front, it would be required to solve an infinite number of optimization problems by varying slightly the objective function for the scalarization method for example. Fortunately, with this method, the information on the slope of each optimum at the Pareto front can be used to approximate accurately the continuous Pareto curve. The method consists in using a piecewise Hermite interpolation between each Pareto optimal value.

As a reminder, we are working on the feasible objective space. Let us define the pair (E_k, F_k) corresponding to an optimal objective value obtained by minimizing the scalarized objective with weight vector $\lambda_k = [\pi w_k, 1 - w_k]$ and $w_k \in [0, 1]$. In addition, from section 1.3.1, we have shown that the gradient at (E_k, F_k) corresponds to $\nabla_k = -\frac{\lambda_{k,1}}{\lambda_{k,2}}$. We consider that the weights are ordered such that $\lambda_0 = (0, 1)$, $\lambda_m = (\pi, 0)$ and supposing that $m+1$ Pareto optimal points are used.

Using the definition of the piecewise Hermite approximation from [14], we have for $e \in [E_k, E_{k-1}]$:

$$H(e) = c_0 + c_1(e - E_k) + c_2(e - E_k)^2 + c_3(e - E_k)^3 \quad (1.6.1)$$

where

$$\begin{aligned} c_0 &= F_k \\ c_1 &= \nabla_k \\ c_2 &= 3 \frac{F_{k-1} - F_k}{h_k^2} - \frac{\nabla_{k-1} + 2\nabla_k}{h_k} \\ c_3 &= \frac{\nabla_{k-1} + \nabla_k}{h_k^2} - 2 \frac{F_{k-1} - F_k}{h_k^3} \end{aligned}$$

with $h_k = E_{k-1} - E_k$. This definition cannot be used in its exact form in the problem considered here. Indeed, the gradient on point (E_m, F_m) is infinite and causes troubles in the approximation. In order to solve this, we consider that the slope has the same growth (called Δ) as the previous one. In other words, we have:

$$\begin{aligned} \Delta &= \frac{F_{m-1} - F_m}{E_{m-1} - E_m} - \nabla_{m-1} \\ \nabla_m &= \frac{F_{m-1} - F_m}{E_{m-1} - E_m} + \Delta \\ &= 2 \frac{F_{m-1} - F_m}{E_{m-1} - E_m} - \nabla_{m-1}. \end{aligned}$$

On Figure 1.6.1 is illustrated the piecewise Hermite approximation for the 3 test cases using 5 points. It returns a smooth function passing through the Pareto optimal values.

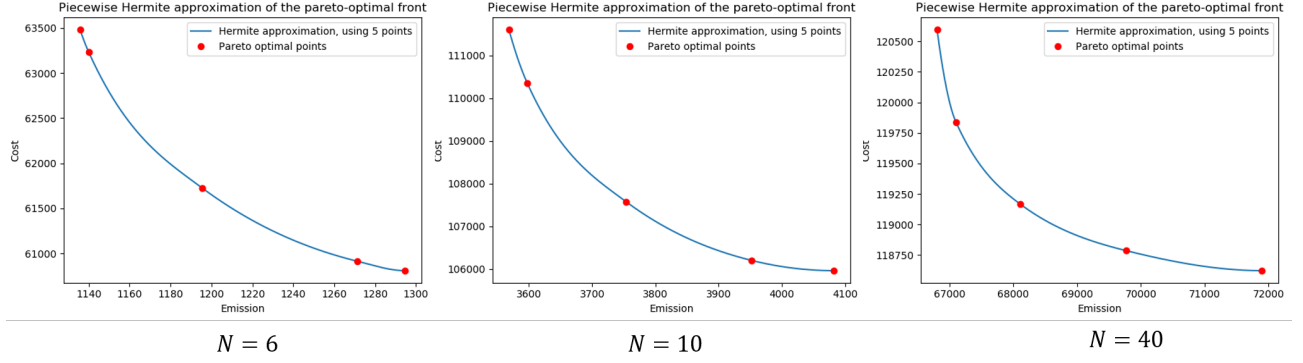


Figure 1.6.1: Piecewise Hermite approximation of the Pareto optimal front.

1.6.1 Error analysis

Let us analyze the accuracy of this approximation comparing to the continuous Pareto front. This reference function is approximated by the region generated by 100 equally spaced weights with the objective given in (1.3.3). By comparing this reference function with the Hermite approximation created by 5 Pareto optimal values, we see from Figure 1.6.2 that the approximation seems to be quite accurate.

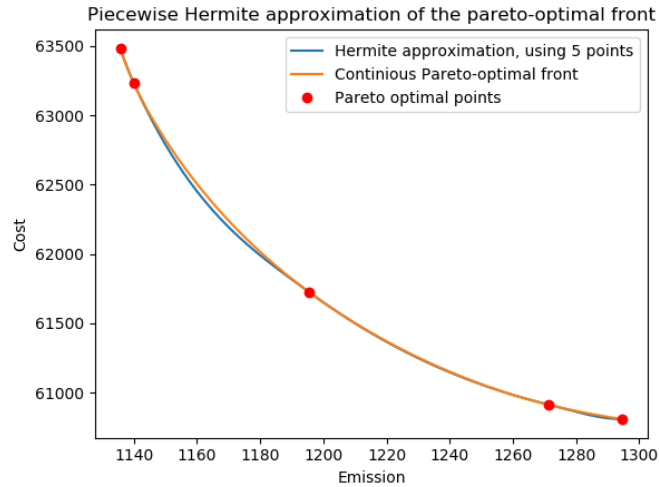


Figure 1.6.2: Analysis of Hermite using 5 points for 6 generators.

In order to measure the generated error on the Hermite interpolation, the infinite norm, or the maximum distance between the two curves is computed. This is an arbitrary choice of quantifying the error but reveals the worst case of using the Hermite approximation. Since the work is done on a parametric graph working with different units, it has more sense to measure the difference following each objective value. Therefore, the worst possible error following one objective, is obtained by comparing two points with same value of the other objective.

The maximal error on the fuel costs is given by:

$$e_F = \max_{1 \leq i \leq 100} (|H(E_i) - F_i|)$$

With $H(E)$ corresponding to the Hermite interpolation and (E_i, F_i) to the i^{th} element of the continuous front. In the example of Figure 1.6.2 using the Hermite approximation with 5 points for the 6-generator test case, the maximum error made on the fuel costs is 54.13 \$ and represents a relative error of less than 0.1%.

The error can also be expressed in terms of lb of total emissions, we use then:

$$e_E = \max_{1 \leq i \leq 100} \left(|H^{-1}(F_i) - E_i| \right)$$

We find here that the maximum error on the emissions is 2.96 lb that corresponds to a relative error of 0.25%.

We conclude therefore that this approximation is accurate, even by using few data points. The Hermite interpolation enables to obtain quickly a reliable estimation of the continuous Pareto front while requiring less Pareto optimal values.

Chapter 2

Dynamic problem

The previous chapter introduced the basic static problem and the methods for obtaining the Pareto optimal front. This chapter relates the possible extensions of the static model by converting it into a dynamic model. The dynamic model takes into consideration the fluctuating demand in fossil fueled energy. Therefore, the problem has to consider a set of different demands related to each time interval and some additional constraints have to be considered.

The first section aims to present the dynamic model and introduces the ramp rates constraints that are peculiar to the dynamic problem. The other constraints that can be considered in this model are the spinning reserves and are explained in the second section. All these extensions are therefore discussed and analyzed in order to observe the changes of the problem and the resulting POF.

2.1 Dynamic demand and ramp rates

It is known that the power demand varies considerably in time [15]. The electric power plants have therefore to be able to adapt their output quickly to respond to this demand. Unfortunately, it is not physical to raise instantly the power generation as much as it is wanted. Indeed, units are limited by their ramp rates that correspond to the rate at which they can increase or decrease their output. This section describes therefore the model of the dynamic problem and analyzes the solutions that can be obtained for the 6-unit test case.

The dynamic problem considers a fluctuating demand in function of the time. Therefore, it is necessary to divide the problem on several time intervals t and compute a dispatch for each interval. Let us work for a one-day problem and divide the day into 24 intervals of time. The power generation is defined as the variable $p_{i,t}$ corresponding to the output of unit i at time t . In addition, the problem has different demands on each time interval defined as $P_{D,t}$. The dynamic Economic Emission Dispatch problem aims at minimizing the total fuel costs and the total emissions during the whole period of activity.

Let us define the down rate DR_i and the up rate UR_i for each generator i . The ramp rates constraint state that the output cannot change too much between two time steps:

$$-DR_i \leq p_{i,t} - p_{i,t-1} \leq UR_i. \quad (2.1.1)$$

The dynamic problem for a one day production can be written as [16]:

$$\begin{aligned}
 \min_{\mathbf{p}} \mathbf{f}(\mathbf{p}) : & \begin{cases} F(\mathbf{p}) = \sum_{i=1}^N \sum_{t=1}^T (a_i + b_i p_{i,t} + c_i p_{i,t}^2) \\ E(\mathbf{p}) = \sum_{i=1}^N \sum_{t=1}^T (\alpha_i + \beta_i p_{i,t} + \gamma_i p_{i,t}^2) \end{cases} \\
 \text{s.t. } \mathbf{p} \in \mathcal{D} : & \begin{cases} \sum_{i=1}^N p_{i,t} = P_{D,t} & t = 1, \dots, T \\ P_i^{\min} \leq p_{i,t} \leq P_i^{\max} & i = 1, \dots, N, \quad t = 1, \dots, T \\ p_{i,t-1} - DR_i \leq p_{i,t} \leq p_{i,t-1} + UR_i & i = 1, \dots, N, \quad t = 1, \dots, T. \end{cases}
 \end{aligned} \tag{2.1.2}$$

As for the static problem, the scalarization technique can also be applied to obtain the POF since the problem stays convex.

This problem is much more realistic because the power plants are most frequently limited by their ramp rates due to drastic changes of demands at different moments of the day. Unfortunately, the literature does not provide data on dynamic demands for all of the studied test cases. We assume therefore to work with a demand function that has the same shape for every test-case. In order to design a demand curve for fossil fueled energy, the approach is based on the famous duck curve of California (see [17]). This demand curve has indeed the shape of a duck due to the balance between the peaks of the demand and the production of renewable energies, able to cover a major part of the total demand in energy in the middle of the day. Note that this shape is not applicable for every region. However, this demand curve is very interesting because in addition to be relevant for the Californian market, the model presents drastic variations making the problem very limited by its ramp rate constraints. Moreover, investments on renewable energies are considerably increasing whole around the world (see [18]) and they will certainly increase their influence on the electric fossil fueled demand in the future. The Californian duck curve is indeed already strongly impacted by the use of solar energies and can describe the future trend of demand curves of other regions.

The production dispatch is displayed on Figure 2.1.1 for the Economic Dispatch and the Emission Dispatch on the dynamic EED. The one-day production is divided into 24 time periods. Recall that the Economic Dispatch consists in lowering the fuel costs without any consideration of the emissions and the Emission Dispatch problem minimizes on the other hand only the emission function. The ramp rate constraints are indeed active when the demand varies drastically, i.e. from 8 to 10 a.m. and from 3 to 6 p.m., which consists in the times of the day when solar power production rises and falls, respectively.

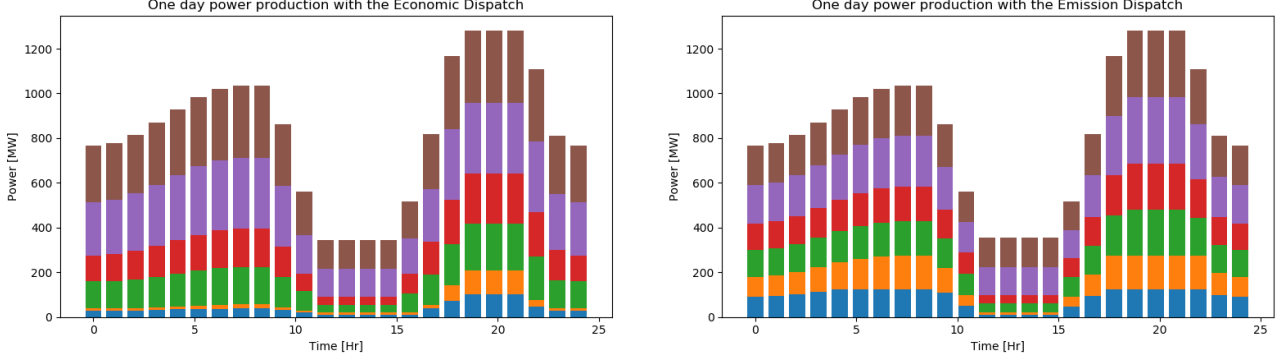


Figure 2.1.1: 24h operation for the Economic (left) and Emission (right) Dispatch with $N=6$.

From this figure, it can be observed that for the 6-unit test case, the generators associated with the blue and orange colour have a low emission function but at a high cost. For this test system, the emissions can be lowered by roughly 2400 lb of pollutants or a diminution of 13.3% from the economic dispatch solution of a one-day production. This leads to an increase of 42.4k \$ corresponding to a raise of 4.1% in fuel costs between the two problems.

2.2 Reserve Constraints

This section describes additional constraints associated with the spinning reserves requirements. The impact on the Pareto optimums and their importance in the problem are discussed in this part of the document.

The transmission system operator (TSO) that is responsible for meeting and transmitting the total power production with the demand of the customers, has to predict in advance the daily demand in energy. Therefore, small unanticipated fluctuations in demand could arise but the TSO has to be able to manage them without interrupting service to clients. They require therefore that the system holds a certain capacity to respond to these variations. The operating power plants have to withhold together a total amount of reserves that is called the spinning reserves [19]. Since each unit can either be limited by its generation limits or by the up ramp rates, the spinning reserve requirements consists in avoiding that all generators are too limited at the same moment.

These spinning reserves are based on the amount of demand that is required at the current time period t . It can be reasonably supposed that the fluctuations during each time interval does not exceed 5% of the current demand. In order to be able to deliver the disturbances within 10 minutes, some reserve constraints can be added for responding to these variations. Here, it is supposed that the variations does not exceed one-third of the maximal variation for one hour. We have therefore that $SR_t = 0.05P_{D,t}$ and $SR'_t = \frac{SR_t}{3}$. The constraints are given in [20]:

$$\bullet \sum_{i=1}^N \min(P_i^{max} - p_{i,t}, UR_i) \geq SR_t$$

$$\bullet \sum_{i=1}^N \min \left(P_i^{max} - p_{i,t}, \frac{UR_i}{6} \right) \geq SR'_t.$$

These constraints can be implemented by using the `min_()` command on **Gurobi**. However, the comparison of the Pareto front between the dynamic demand with and without reserves, illustrated on Figure 4.2.1, is not significant. The illustration is based on the 10-unit test system and this example presents a difference in cost of 0.075% for the Economic Dispatch and 0.033% difference in emission for the Emission Dispatch. It is also possible to note that the emissions can be considerably reduced from the Economic Dispatch solution. Indeed, the emissions can be reduced of at most 8% while only modifying the production dispatch. Note that this extreme result leads to a high increase in costs and is therefore not the most interesting solution. Recall from section 1.2.2 that the Kalai-Smorodinsky solution is a good method for obtaining a compromise between fuel cost reduction and the environmental impact. As a result, the KS solution provides a reduction of 5.8% of pollutant emissions with an increase in costs of 1.2% comparing to the Economic Dispatch. This result seems indeed to be very attractive for real-life power systems.

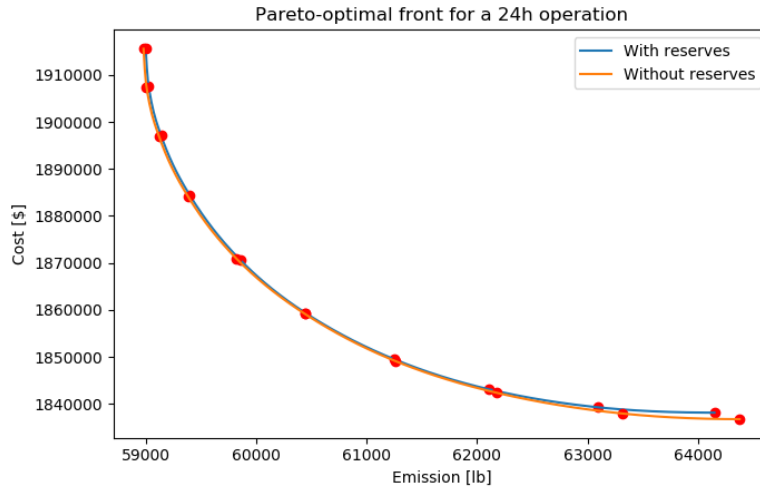


Figure 2.2.1: Influence of the reserve constraints for a 24h operation with $N=10$. Note that the red dots are the computed Pareto optimal values and are linked with the Hermite approximation.

As a conclusion, the conversion from static to dynamic problem does not affect the nature of the problem as it remains a convex Quadratic Programming (QP) problem. However, it computes a more realistic visualisation of the power production. One can therefore note the importance of the EED comparing to the classic ED problem. Indeed, even if the observed test-case is fictive, it reveals that it is possible to modify the optimal dispatch in order to have an impacting reduction of the emissions for only a slight increase in costs.

On another side, it is possible to have a more accurate model by reducing the time steps and deal with more precise demands. However, this would lead to a huge increase in computational time due to the increase in the number of variables. A final remark is therefore that even if the dynamic model represents more genuine figures, it requires much more time for being solved than the static problem. This chapter motivates therefore to drop the dynamic model for these reasons and to work with the static problem as it will be done in Chapter 4.

Chapter 3

Nonlinear Emission objective

This chapter describes a more accurate approximation of the total pollutant emissions of each power plant. Indeed, the two first chapters are based on a quadratic emission function in order to work with a Quadratic Program when the objective function is scalarized. This chapter explains therefore the impact of adding the nonlinear term in the emission function.

3.1 Exponential term in the emissions

From [21], it is shown that a better estimation on the emission rejection for each unit can be made by modelling this function as nonlinear. Indeed, an exponential factor influences this rate and the problem is no longer a QP problem.

A better expression of the emission function for one unit is given by:

$$E_i(\mathbf{p}) = (\alpha_i + \beta_i p_i + \gamma_i p_i^2 + \eta_i e^{\delta_i p_i}) \quad (3.1.1)$$

with $\boldsymbol{\eta}, \boldsymbol{\delta} \in \mathbb{R}_+^N$. By considering problem (2.1.2) with the additional constraints on the spinning reserves described in section 2.2, only the emission objective is modified and corresponds to :

$$E(\mathbf{p}) = \sum_{i=1}^N \sum_{t=1}^T (\alpha_i + \beta_i p_{i,t} + \gamma_i p_{i,t}^2 + \eta_i e^{\delta_i p_{i,t}}). \quad (3.1.2)$$

Since the problem remains convex, the scalarization method is still suited for computing the Pareto optimal front. However, **Gurobi** has to consider a piecewise linear function for approximating the exponential function [22]. This can be done by adding a constraint for linearizing the exponential function by using the `addGenConstrExp()` command. The optimum will therefore not be exact anymore. However, given that it returns a feasible solution with a tolerance of 10^{-3} on the approximation of the exponential, the result can be accepted with confidence.

The comparison of the Pareto fronts obtained with the quadratic emission function and the nonlinear function is illustrated on Figure 4.2.1 for the 10-unit test case. Roughly, it can be seen as a translation of the curve to the right because only the emission function is impacted. It is indeed not an exact translation because the Pareto optimal solutions differ from the quadratic problem. With the additional exponential factor in the emission objective, this latter value raises by 1.67% or approximately 1000 *lb* for the Environmental Dispatch problem (top left on

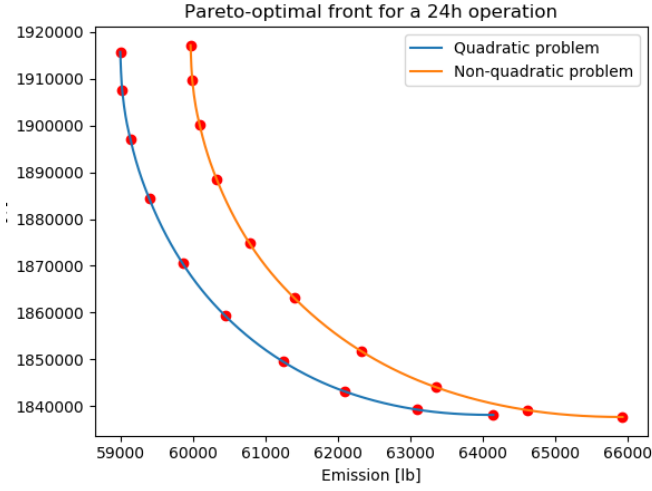


Figure 3.1.1: Influence of the Pareto optimal front by adding the exponential term, $N=10$.

the curve) and increases by 2.68% or 1700 *lb* for the Economic Dispatch (bottom right).

As described in this chapter, the exponential term in the emission objective does not change the convexity of the problem but increases the complexity for the solver. The conversion from a QP to a Nonlinear Program (NLP) results to an increase in computational time. However, it modifies the estimation on the emissions from 1% to 3%. Since our work consists mainly in analyzing the impact of the emissions, this factor has therefore a significant importance.

Chapter 4

Transmission losses

Until now, the previous chapters neglected the losses in the system but it is observed that they impact largely the solutions. The following chapter studies therefore the augmented problem with the transmission losses. The EED becomes non-convex and leads to difficulties in the resolution. Indeed, non-convex problems are not guaranteed to have a unique solution and traditional solvers does not ensure to provide convergence for example. Therefore, several tests and analyses are done in order to solve this more complex problem.

The first section presents the model for the transmission losses considered here, namely the Kron's coefficients. Then, a preliminary analysis on the behaviour of the Pareto optimal front is realised to give some insights on the challenges related to non-convex multi-objective problems. Section 4.3 describes two iterative methods in order to solve the non-convex problem.

Then, an enlarged study of the feasible domain is achieved in Section 4.4 in order to characterize the feasible set and the resulting solutions. Since it is possible to treat the complications related to non-convex programs by relaxing the problem, some approaches on relaxations are proposed in the next section based on this analysis. Section 4.6 presents and formalises three relaxations and studies their validity. The next section aims at comparing and measuring the performances of each relaxation.

Finally, the two last sections explain different powerful solvers for the convex relaxation problem and analyze their efficiency. Section 4.8 describes three first order methods and the last section focuses on second order methods.

4.1 Non-convex problem

An important parameter for modelling this problem consists in the losses during the transmission. Indeed, a non negligible amount of power is lost during the power transmission and it must therefore be taken into consideration. This first section describes the upgraded model and evaluates the difficulties associated to multi-objective non-convex programming.

The transmission losses play a role in the demand constraint for every period of time, it becomes now:

$$\sum_{i=1}^N p_{i,t} = P_{D,t} + p_{L,t} \quad t = 1, \dots, T$$

The transmission losses $p_{L,t}$ can be expressed as a quadratic function of the power output of each unit and is given in [10] :

$$p_{L,t} = \sum_{i=1}^N \sum_{j=1}^N p_{i,t} B_{ij} p_{j,t}$$

where the elements B_{ij} are known as the loss coefficients. This document supposes to work only with quadratic losses and by ruling out the possible linear losses that may occur. This is because the linear losses does not affect fundamentally the complexity of the problem and also because this thesis bases its problem formulation mainly on [9] and [10] that do not provide data on these coefficients. We also make the assumption that the loss matrix B is symmetric positive definite.

This quadratic equality constraint makes the problem non-convex and could therefore create troubles due to local minima. Moreover, non-convex multi-objective problems cannot be solved by using the scalarization of the objective in order to obtain all Pareto optimal values. Indeed, Figure 4.1.1 illustrates as an example a set of achievable objective values for a non-convex problem. The optimal front is not guaranteed to be convex meaning that certain Pareto optimal values cannot be reached with the scalarization method. On the other hand, the ϵ -constraint method is still able to retrieve all Pareto optimal values.

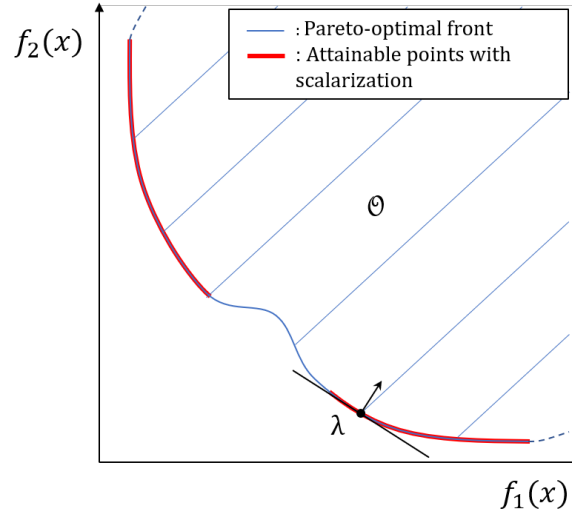


Figure 4.1.1: Achievable objective values for a non-convex problem.

4.2 Pareto front analysis

This section aims at giving some insights with respect to the challenges linked to non-convex multi-objective problems. Indeed, the feasible set does not seem to contain much irregularities in spite of its non-convex definition. The two methods for obtaining the Pareto front are used and described in order to reveal if the Pareto optimal front remains convex.

4.2.1 Scalarization method

First of all, the scalarization of the objective can be applied to obtain, at least partially, the Pareto optimal front. A first idea of the representation of this front can be build up from this analysis.

Gurobi supports non-convex quadratic constraints but it leads to a huge increase in computational time. An estimation of the front is realised for the 10-unit test system and is represented

on Figure 4.2.1. Note that the demand function had to be scaled down in order to keep a feasible problem. We observe the influence of the losses by comparing the Pareto optimal values between the convex and non-convex problem. In order to measure the offset between the two resulting curves, we observe an increase in fuel costs of approximately 50k \$ or 3% close to the Economic Dispatch and an increase in emissions of approximately 2200 lb corresponding to a relative increase of 4% near the Emission Dispatch.

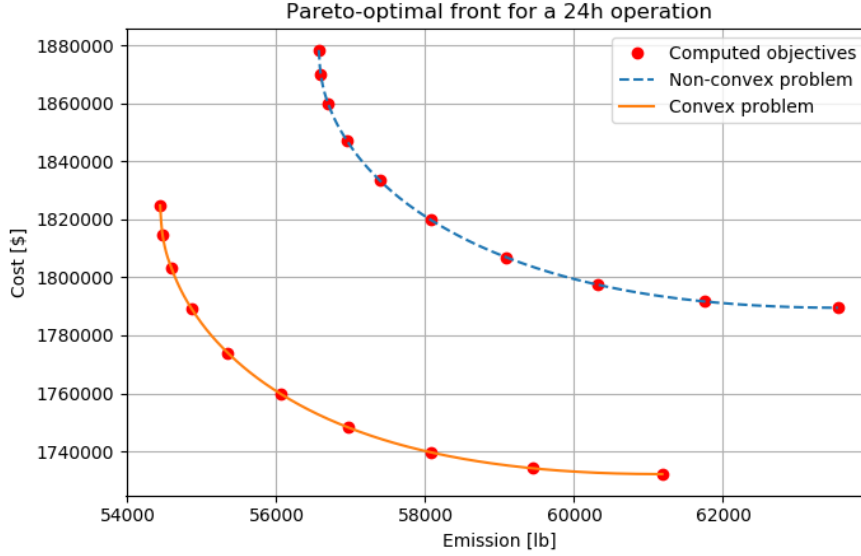


Figure 4.2.1: Comparison of the Pareto optimal fronts of the problem with transmission losses for the 10-unit test system. The red dots are computed from the Pareto optimal values, the lines consists in the Hermite approximation.

This first analysis does not highlight a non-convex region on the Pareto front because the slope evolves regularly and the spacing between the Pareto values are equivalent to those of the convex problem. It seems that the non-convexity due to the transmission losses does not affect the convexity of the POF by observing the smooth approximation between each points. However, based on such few points, it gives us only a general idea and no formal proof.

4.2.2 ϵ -constraint method

Remind that this technique consists in minimizing a single objective while constraining the other objective with different limits. This method is more appropriate for non-convex problems since it can reach all Pareto optimal values regardless of the convexity of the front [23]. Therefore, let us study the solutions that are obtained with this method in order to reveal possible irregularities on the Pareto front.

The ϵ -constraint problem on the emissions can be written as:

$$\begin{aligned}
 \min F(\mathbf{p}) &= \sum_{i=1}^N \sum_{t=1}^T (a_i + b_i p_{i,t} + c_i p_{i,t}^2) & (4.2.1) \\
 \text{s.t. } E(\mathbf{p}) &= \sum_{i=1}^N \sum_{t=1}^T (\alpha_i + \beta_i p_{i,t} + \gamma_i p_{i,t}^2 + \eta_i e^{\delta_i p_i}) \leq E_i & E_{\min} \leq E_i \leq E_{\max} \\
 \sum_{i=1}^N p_{i,t} &= P_{D,t} + p_{L,t} & t = 1, \dots, T \\
 p_{L,t} &= \sum_{i=1}^N \sum_{j=1}^N p_{i,t} B_{ij} p_{j,t} & t = 1, \dots, T \\
 P_i^{\min} &\leq p_{i,t} \leq P_i^{\max} & i = 1, \dots, N, t = 1, \dots, T \\
 p_{i,t-1} - DR_i &\leq p_{i,t} \leq p_{i,t-1} + UR_i & t = 1, \dots, T \\
 \sum_{i=1}^N \min(P_i^{\max} - p_{i,t}, UR_i) &\geq SR_t & t = 1, \dots, T \\
 \sum_{i=1}^N \min\left(P_i^{\max} - p_{i,t}, \frac{UR_i}{6}\right) &\geq SR'_t & t = 1, \dots, T
 \end{aligned}$$

The ϵ -constraint problem on the costs is an analog problem except that the objective is not quadratic. By trying the two methods with a set of limits in the emissions or costs, the graph in Figure 4.2.2 can be obtained by plotting the results for the different methods. The green crosses correspond to some objective values with fixed emission, the purple crosses are obtained by limiting the fuel costs with different values and the red dots are the results with the scalarization method. The dashed blue line is the hypothetical Pareto front constructed from the results of the scalarization method. All these points lie on this Hermite approximation which strengthen the hypothesis that the POF remains convex.

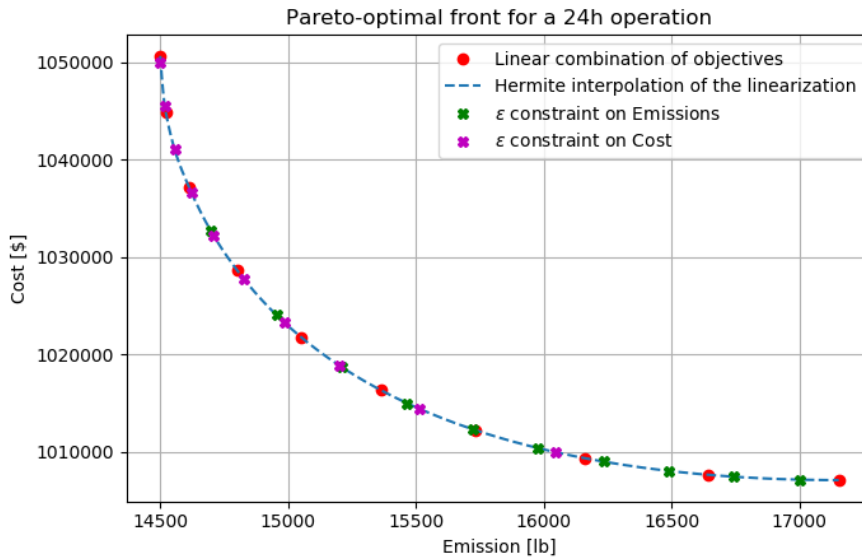


Figure 4.2.2: Pareto optimal values obtained from different methods for the 6-unit test system.

4.3 Gradient based methods on non-convex programming

Solving a non-convex problem with a nonlinear objective is computationally expensive to solve. In fact, it is observed that it takes approximately 50 minutes for computing 10 Pareto optimal values on the 10-unit test system with the direct implementation in Gurobi. This section presents therefore traditional optimization algorithms based on the gradient descent in order to try their effectiveness on the non-convex problem.

The Constrained Gradient Descent algorithm and the Accelerated Gradient Method are studied in this section. Unfortunately, these two methods does not guarantee the convergence to the global minimum for non-convex problems. The purpose of this section is therefore to verify if these first order solvers work and if they converge to the global minimum. In addition, their performances are also studied in order to characterize their efficiency.

Algorithms based on the gradient descent can indeed be analyzed and their performances can be studied for this non-convex problem. As shown in Chapter 2, the dynamic problem does not complicate the type of problem - a non-convex program with a nonlinear objective function - but increases largely the computation time. Hence, the analyses are based on the static problem in order to lighten the problem. Note that for the rest of this report, the static problem will be considered for sake of simplicity. The solvers that are studied here focus on the scalarized problem because it allows a more efficient implementation. Indeed, nonlinear functions are easier to handle in the objective than in the constraints as required by the ϵ -constraint method. Moreover, we will see in section 4.4.3 that this method is guaranteed to reach all Pareto optimal solutions given certain conditions on the parameters. The following problem is therefore considered:

$$\begin{aligned}
 \min \quad & f(\mathbf{p}) = \pi w_1 E(\mathbf{p}) + w_2 F(\mathbf{p}) \\
 \text{s.t.} \quad & \mathbf{p} \in \mathcal{D} : \begin{cases} \sum_{i=1}^N p_i - p_L = P_D \\ p_L = \sum_{i=1}^N \sum_{j=1}^N p_i B_{ij} p_j \\ P_i^{\min} \leq p_i \leq P_i^{\max} \quad i = 1, \dots, N. \end{cases}
 \end{aligned} \tag{4.3.1}$$

4.3.1 Constrained gradient descent

Firstly, let us analyze the objective function and consider, without loss of generality, that $w_1 = w_2 = 1$. The constrained gradient method is proved to be effective for convex problems [24]. Since we are working with a convex objective but in a non-convex set $\mathcal{D}$, it is not guaranteed that the method converges to a global minimum. The Lipschitz constant may be useful for using the gradient descent algorithm. Since f is continuous and twice differentiable, we know that its gradient is Lipschitz continuous with constant L given that $\|\nabla^2 f\| \leq L$.

In other words, we have $L = \max(\lambda(\nabla^2 f))$. Let us compute the Lipschitz constant for the function f in (4.3.1):

$$\begin{aligned}
 \lambda_i &= \pi \left[\beta_i + 2\gamma_i p_i + \delta_i^2 \eta_i e^{(\delta_i p_i)} \right] + b_i + 2c_i p_i. \\
 L &= \max_{1 \leq i \leq N} (\lambda_i) \leq \max_{1 \leq i \leq N} \pi \left[b_i + 2\gamma_i P_i^{\max} + \delta_i^2 \eta_i e^{(\delta_i P_i^{\max})} \right] + b_i + 2c_i P_i^{\max}.
 \end{aligned}$$

We also note that the objective function has the property of being strongly convex. Indeed, a twice differentiable function f is μ -strongly convex if $\|\nabla^2 f\| \geq \mu$. We have:

$$\|\nabla^2 f\| \geq \min_{1 \leq i \leq N} (\lambda_i) \geq \min_{1 \leq i \leq N} \pi \left[b_i + 2\gamma_i P_i^{min} + \delta_i^2 \eta_i e^{(\delta_i P_i^{min})} \right] + b_i + 2c_i P_i^{min} = \mu$$

The gradient descent algorithm goes as follows [24]:

$$\begin{aligned} \mathbf{P}_0 &\in \mathcal{D} \\ \mathbf{P}_{k+1} &= \text{Proj}_{\mathcal{D}}[\mathbf{P}_k - h \cdot \nabla f(\mathbf{P}_k)] \end{aligned} \quad (4.3.2)$$

with $h = \frac{2}{\mu+L}$, initial point $\mathbf{P}_0 = \text{Proj}_{\mathcal{D}}[\mathbf{P}^{min}]$ and the projection function is given by:

$$\text{Proj}_{\mathcal{D}}[\mathbf{P}_k] = \arg \min_{\mathbf{p} \in \mathcal{D}} \|\mathbf{p} - \mathbf{P}_k\|^2$$

However, it is not guaranteed that the Gradient method converges. It is indeed shown on the left of Figure 4.3.1, that the method fails when the strong convexity is used. However, as displayed on the right, the constrained gradient method is able to converge to the global minimum with a smaller step size corresponding to $h = \frac{1}{L}$. It is noted that it reaches the optimum with tolerance 10^{-2} after 28 iterations.

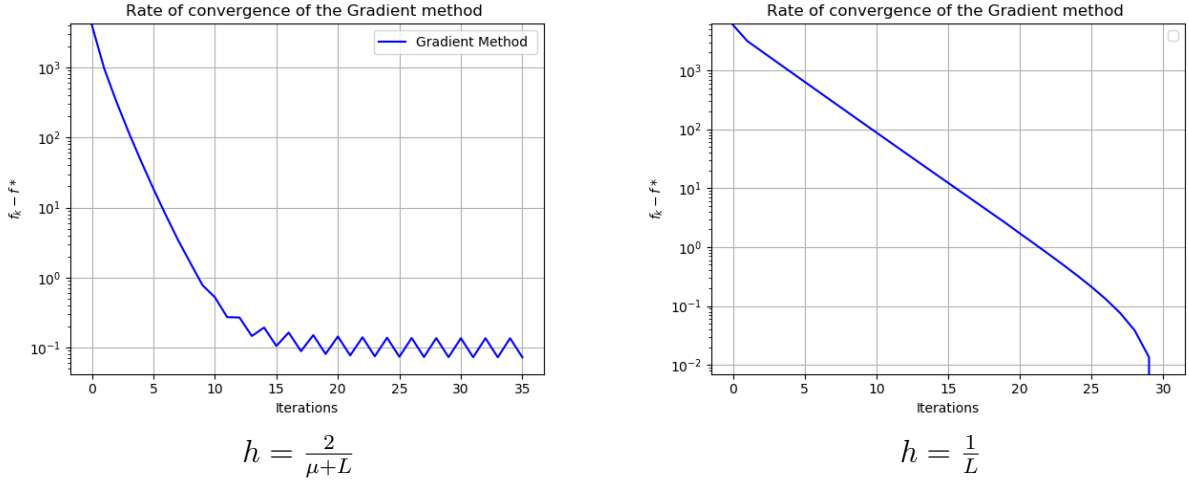


Figure 4.3.1: Rate of convergence for different step sizes on the 10-unit test case.

4.3.2 Accelerated Gradient

It is possible to optimize this algorithm with the accelerated gradient technique. The L -Lipschitz and μ -strongly convex properties on the objective function are also used. The algorithm is given by the following steps [24]:

$$\begin{aligned} \mathbf{y}_0 &= \mathbf{P}_0 \in \mathcal{D} \\ \mathbf{P}_{k+1} &= \text{Proj}_{\mathcal{D}}[\mathbf{y}_k - h \cdot \nabla f(\mathbf{y}_k)] \\ \mathbf{y}_{k+1} &= \mathbf{P}_{k+1} + \beta(\mathbf{P}_{k+1} - \mathbf{P}_k) \end{aligned} \quad (4.3.3)$$

The optimal step is $h = \frac{1}{L}$ and we have $\beta = \frac{\sqrt{L}-\sqrt{\mu}}{\sqrt{L}+\sqrt{\mu}}$. As we can see from Figure 4.3.2, the accelerated gradient converges faster to the solution, taking only 14 iterations for the problem with 10 generators.

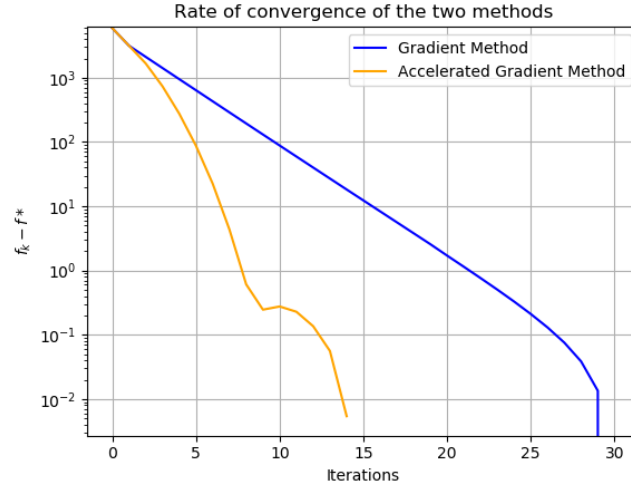


Figure 4.3.2: Comparison of the two methods for the static problem with 10 generators

We can therefore conclude that it is possible that these classic gradient-based methods work on this specific non-convex domain $\mathcal{D}$. From these two examples, the obtained iterates converge to the global minimum in a few iterations and the accelerated gradient performs much better as it is expected. However, even if these first order solvers works on problems with the test cases of 6 and 10 generators, numerical issues are encountered when applying this algorithm to the 40-unit test system.

If we compare the performances of the Accelerated method with the direct implementation of the problem in `Gurobi`, this latter solver remains more powerful. Indeed, it takes about 5 seconds to solve one non-convex problem while the accelerated gradient takes approximately 15 seconds. On the other hand, it is interesting to highlight that this iterative method has to solve 14 quadratic non-convex problems, corresponding to the projections, while the direct implementation has to solve one non-quadratic non-convex problem. We will see indeed in section 4.8 that these algorithms present some potential with the convex relaxation of the problem.

4.4 Analysis of the achievable domain of objectives

Based on the results of sections 4.2 and 4.3, it is possible that this non-convex problem presents some properties that allow to make assumptions on its behaviour. This section is dedicated to the analysis of the feasible domain and attempts to show that the Pareto optimal front is convex. An enlarged analysis is made for a simple 2-generator and 3-generator problem in order to obtain some graphic observations.

4.4.1 2 generator analysis

The major problem of non-convex optimization programming is that it is not guaranteed that it has only global minimums. This section studies therefore the feasible domain for the problem using 2 generators and observes any irregularities on the achievable set of objective values. Remind that we are working with the simplified static problem stated by (4.3.1). The demand

constraint is therefore equal to:

$$\begin{aligned} \sum_{i=1}^N \sum_{j=1}^N p_i B_{ij} p_j - \sum_{i=1}^N p_i + P_D &= 0 \\ \Leftrightarrow \begin{bmatrix} p_1 & p_1 & 1 \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & -\frac{1}{2} \\ B_{21} & B_{22} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & P_D \end{bmatrix} \begin{bmatrix} p_1 \\ p_1 \\ 1 \end{bmatrix} &= 0 \\ \Rightarrow \mathbf{x}^T M \mathbf{x} &= 0 \end{aligned}$$

It can be noticed that this equality constraint is the quadratic equation of a conic section. Since B is symmetric positive definite, the expression describes an ellipse [25].

Let us analyse all points in the domain: they are determined by the equation $\mathbf{x}^T M \mathbf{x} = 0$ and are included in the feasibility box defined as $[\mathbf{P}^{min}, \mathbf{P}^{max}]^2$. The system described with the ellipse can be converted in a parametric equation. This requires to compute the radius of the x- and y-axes, the rotation and the center of the ellipse.

- If we call λ_1 and λ_2 the eigenvalues of matrix B , the radii are:

$$\begin{aligned} r_1 &= \sqrt{\frac{-\det(M)}{\lambda_1 \det(B)}} = \frac{1}{\sqrt{\lambda_1}} \sqrt{\frac{1}{4} \mathbf{e}^T B^{-1} \mathbf{e} - P_D} \\ r_2 &= \sqrt{\frac{-\det(M)}{\lambda_2 \det(B)}} = \frac{1}{\sqrt{\lambda_2}} \sqrt{\frac{1}{4} \mathbf{e}^T B^{-1} \mathbf{e} - P_D} \end{aligned} \quad (4.4.1)$$

With $\mathbf{e}$ being the vector of ones of size 2.

- The rotation matrix is given by the unitary matrix U such that:

$$U^T B U = \Lambda.$$

We have therefore for $\mathbf{v}^i$, the normalized eigenvector of λ_i :

$$U = \begin{bmatrix} v_x^1 & v_x^2 \\ v_y^1 & v_y^2 \end{bmatrix} \quad (4.4.2)$$

- The center is obtained as follows:

$$\begin{bmatrix} x_c \\ y_c \end{bmatrix} = -B^{-1} \begin{bmatrix} M_{13} \\ M_{23} \end{bmatrix} = B^{-1} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}. \quad (4.4.3)$$

We can now define the parametric equation for $\theta \in [0, 2\pi]$:

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} v_x^1 & v_x^2 \\ v_y^1 & v_y^2 \end{bmatrix} \begin{bmatrix} r_1 \sin(\theta) \\ r_2 \cos(\theta) \end{bmatrix} + \begin{bmatrix} x_c \\ y_c \end{bmatrix} \quad (4.4.4)$$

For the problem with 2 generators, the ellipse is illustrated on Figure 4.4.1 for the 2-unit test case. This test case is a simplified version of the 6-unit test case as described in Appendix A.1. It can be noted that it has a very large radius following the two axis due to the small eigenvalues of the transmission loss matrix B . The feasible domain concerns only a small region

of this ellipse since it has to satisfy the box constraints given by the maximal and minimal generation of the units.

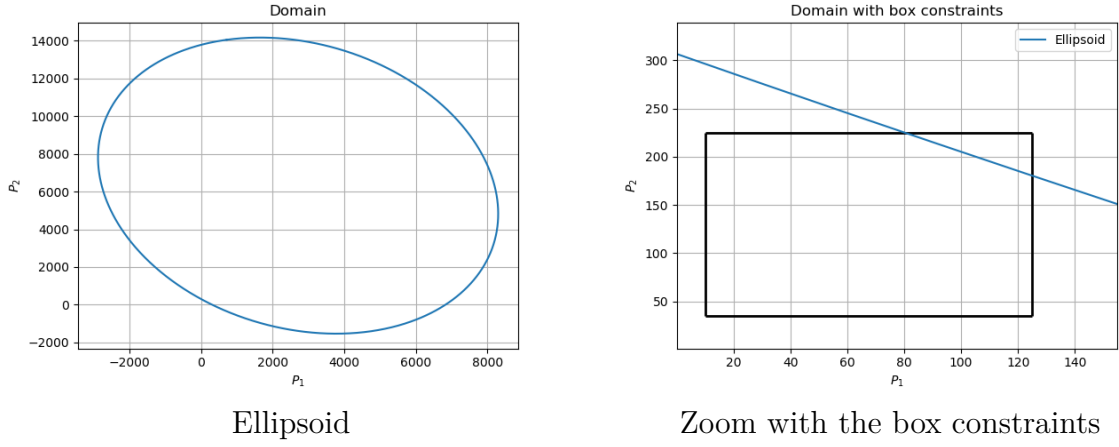


Figure 4.4.1: Feasibility Domain.

The achievable set of objective values can be obtained by browsing the elliptic curve inside the box and by computing their corresponding costs and emissions. The set, given in Figure 4.4.2, corresponds to a single curve, and also describes the Pareto optimal front. It means that every feasible solution is therefore Pareto optimal and the POF remains convex.

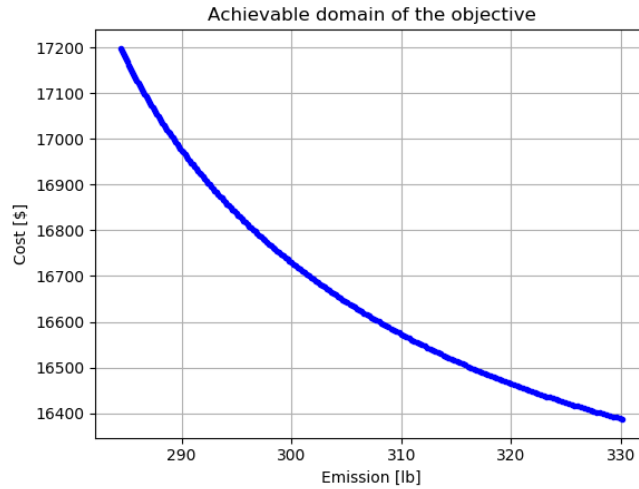


Figure 4.4.2: Achievable objective values.

4.4.2 3 generator analysis

In order to have a more general idea of the behaviour of this problem for multiple generators, the analysis is illustrated on a problem with 3 generators. The 3-unit test case is described in Appendix A.2. Instead of working with a 2-dimensional ellipse, the demand constraint can be represented as an ellipsoid. Its parameters are obtained with the same formulas as in the previous subsection and by extending the problem into a 3-dimensional problem.

This results in the following parametric equation for $\theta \in [0, 2\pi]$ and $\phi \in [0, 2\pi]$:

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = U \begin{bmatrix} r_1 \cos(\phi) \sin(\theta) \\ r_2 \cos(\phi) \cos(\theta) \\ r_3 \sin(\phi) \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \quad (4.4.5)$$

By analyzing the feasibility region defined by the maximum and minimum production per unit, we are looking at the part of the ellipsoid that is contained inside the feasibility box. On the left of Figure 4.4.3 is displayed the intersection of the ellipsoid with the feasibility box. The domain of definition is shown in red on the figure. On the right is displayed the achievable domain of the objectives. The optimal front remains convex as for the 2-dimensional problem.

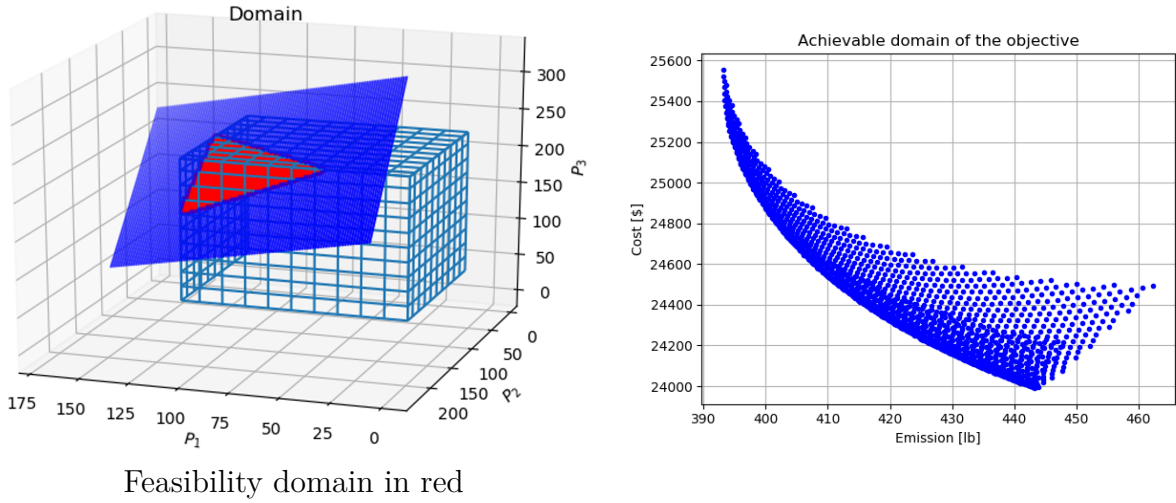


Figure 4.4.3: Analysis for a 3 unit problem.

4.4.3 Generalisation

From the two latter subsections, it is tempting to generalize that the Pareto optimal front is always convex as it is the case for the test case with 2 and 3 generators. Unfortunately, this assumption is difficult to demonstrate on a general problem.

However, it will be explained in section 4.6.1 that it is possible to obtain an exact convex relaxation of the feasible domain given some conditions on the parameters. Therefore, when this relaxation is exact, this problem is equivalent with the non-convex problem and the corresponding POF is necessarily convex.

4.5 Approach to relax the non-convex demand

As it is observed in the previous section, the surface defined by the ellipsoid is almost characterized as a flat plane. Indeed, the feasibility box is much smaller in comparison with the size of the ellipsoid. It is therefore interesting to build a linear approximation and to develop some relaxations on this non-convex demand constraint. This section gives a primary analysis on possible relaxations based on the examples of 2 and 3 generators in the system. A proper analysis of these approaches will be studied in Section 4.6.

4.5.1 (Hyper)plane constructed with N extreme points

From the previous study, it can be noticed that the feasible box defined by $[\mathbf{P}^{min}, \mathbf{P}^{max}]^N$ is located on the bottom left of the ellipsoid. Therefore, the hypercurve that is inside this box can be considered as convex and decreasing. This is an important feature because this may help for constructing a linear upper bound. This section describes indeed how to obtain a linear equation from the quadratic demand equality. A first approach is done for a 3-unit problem and the method is then explained for a general problem with N generators.

Firstly, let us consider working with the 3-unit test system. The hypercurve defined with the demand intercepts the feasibility box defined by $[\mathbf{P}^{min}, \mathbf{P}^{max}]^3$ on 3 edges. Therefore these 3 extreme points that are located at the edges of the feasibility box are sufficient for obtaining a plane that is superior to all feasible points. These extreme points are described as:

$$\begin{aligned} \mathbf{P}^1 &= (x, P_2^{max}, P_3^{max}) & x : \begin{bmatrix} x \\ P_2^{max} \\ P_3^{max} \\ 1 \end{bmatrix}^T M \begin{bmatrix} x \\ P_2^{max} \\ P_3^{max} \\ 1 \end{bmatrix} &= 0 \\ \mathbf{P}^2 &= (P_1^{max}, y, P_3^{max}) & y : \begin{bmatrix} P_1^{max} \\ y \\ P_3^{max} \\ 1 \end{bmatrix}^T M \begin{bmatrix} P_1^{max} \\ y \\ P_3^{max} \\ 1 \end{bmatrix} &= 0 \\ \mathbf{P}^3 &= (P_1^{max}, P_2^{max}, z) & z : \begin{bmatrix} P_1^{max} \\ P_2^{max} \\ z \\ 1 \end{bmatrix}^T M \begin{bmatrix} P_1^{max} \\ P_2^{max} \\ z \\ 1 \end{bmatrix} &= 0 \end{aligned}$$

The unknowns x, y and z can be found by solving a quadratic equation. Once these 3 points $\mathbf{P}^1, \mathbf{P}^2, \mathbf{P}^3$ are computed, we can obtain 2 director vectors $\overrightarrow{\mathbf{P}^1 \mathbf{P}^2}$ and $\overrightarrow{\mathbf{P}^1 \mathbf{P}^3}$ and obtain the normal vector to the plane by solving the following linear equation:

$$\begin{bmatrix} \overrightarrow{\mathbf{P}^1 \mathbf{P}^2} \\ \overrightarrow{\mathbf{P}^1 \mathbf{P}^3} \\ 1 \ 0 \ 0 \end{bmatrix} \mathbf{n} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

The equation of the resulting plane is given by:

$$\sum_{i=1}^N n_i p_i - \underbrace{\sum_{i=1}^N P_i^1 n_i}_{=k} = 0.$$

Let us analyze the error that is created between the quadratic equation and the linear equality that is described above. The error is given by the distance between these two surfaces. Figure 4.5.1 illustrates the linear approximation on the left and the error at each point on the right. It results that the maximum error of this approximation is less than 0.16 MW corresponding to a relative error of $3.73 \times 10^{-2} \%$.

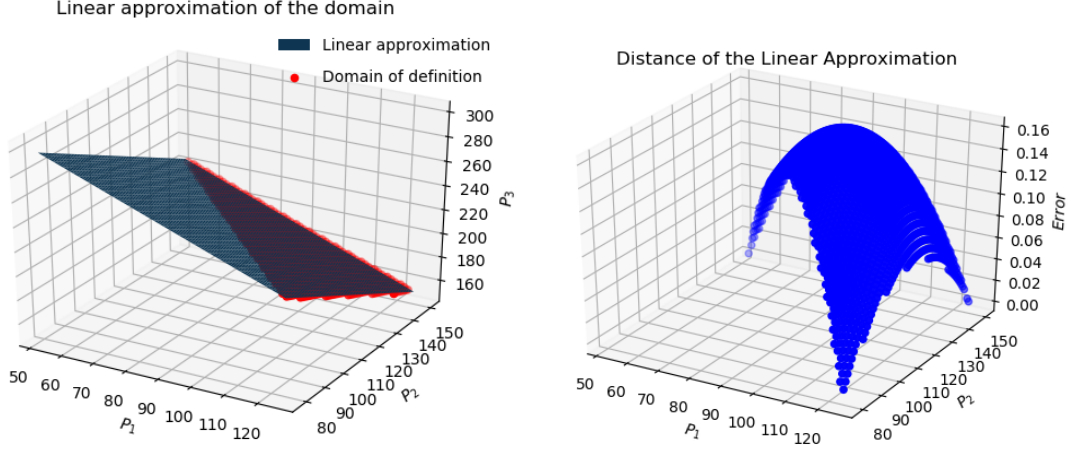


Figure 4.5.1: Analysis of the linear approximation.

From this study, a general method can be constructed for obtaining a linear upper bound on a problem of N power plants. The first step consists in computing the N extreme points located on the edges of the feasible box $[\mathbf{P}^{min}, \mathbf{P}^{max}]^N$. Since there are $N2^{N-1}$ candidates, or edges, the method consists in exploring them until it has found N points intercepting the demand equality. The detailed algorithm for obtaining these points will be described in section 4.6.2.

Once we have found the N points, they all have to satisfy the linear constraint. In order to achieve this, the normal vector $\mathbf{n}$ has to be computed and corresponds to the coefficients of the variables in this constraint. If we call $\mathbf{P}^j$ all extreme points, we have to solve the following system:

$$\sum_{i=1}^N n_i P_i^j + k = 0 \quad j = 1, \dots, N. \quad (4.5.1)$$

It can be converted in the following linear system:

$$\begin{bmatrix} P_1^1 & P_2^1 & \dots & P_N^1 & 1 \\ P_1^2 & P_2^2 & \dots & P_N^2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ P_1^N & P_2^N & \dots & P_N^N & 1 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_N \\ k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Note that there are $N + 1$ variables for N equations, hence the constraint that $n_1 = 1$ is added. By solving this system, vector $\mathbf{n}$ and scalar k are obtained and the linear upper bound constraint can be formulated as:

$$\sum_{i=1}^N n_i p_i + k \leq 0. \quad (4.5.2)$$

4.5.2 Feasible set: convex relaxation

Since a linear constraint that limits superiorly all feasible points is established, a convex relaxation of the quadratic equality can be obtained. By considering the set of points included above the quadratic demand constraint and under the linear upper bound, we obtain a convex relaxation of the problem. The relaxed problem has therefore the following form:

$$\begin{aligned}
 \min_{\mathbf{p}} \quad & f(\mathbf{p}) \\
 \text{s.t.} \quad & \mathbf{p}^T B \mathbf{p} - \mathbf{e}^T \mathbf{p} + P_D \leq 0 \\
 & \mathbf{P}^{min} \leq \mathbf{p} \leq \mathbf{P}^{max} \\
 & \mathbf{n}^T \mathbf{p} + k \leq 0
 \end{aligned} \tag{4.5.3}$$

with $\mathbf{e}$ being the vector of ones of size N . Such a feasible domain can be sketched for a problem with 2 power plants as illustrated on the left on Figure 4.5.2. On the right in the figure is displayed the true feasible set that is contained between the blue and orange curve. The Pareto optimal values computed for some weights on the objectives all tighten the demand constraint. The relaxation is therefore exact on this problem.

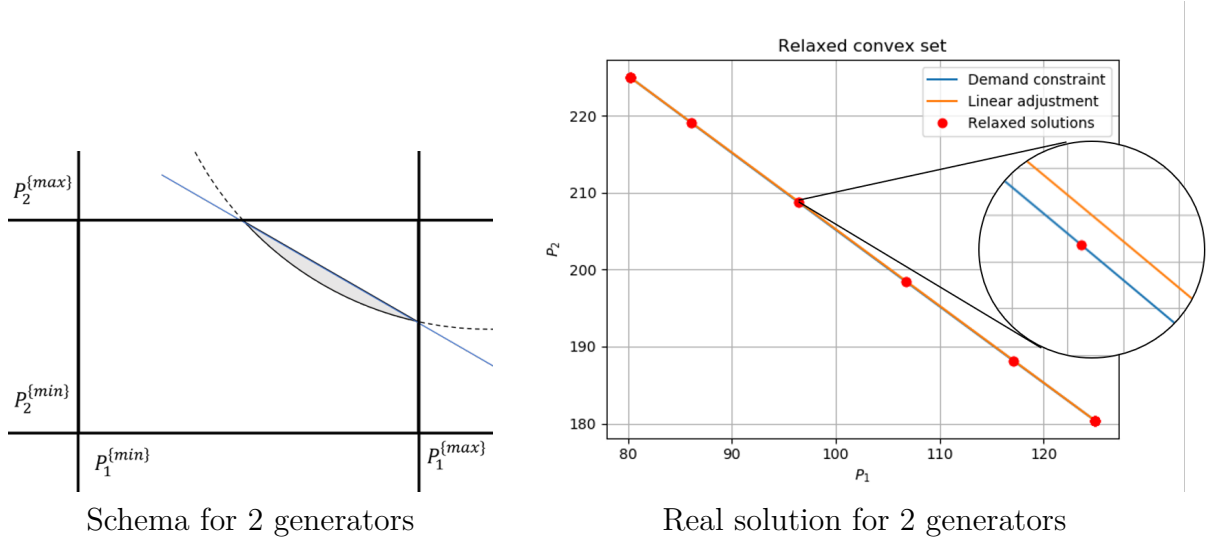


Figure 4.5.2: Relaxed Convex set with linear upper bound.

It is thus not required to add the linear upper bound to relax the problem since this constraint is not active on any Pareto optimal value. This means that the feasible set considered as schematized on Figure 4.5.3 keeps the same solutions. It has indeed a physical sense because replacing the equality by an inequality simply consists in overestimating the losses. The proof of the exactitude of this relaxation will be demonstrated in detail in section 4.6.1.

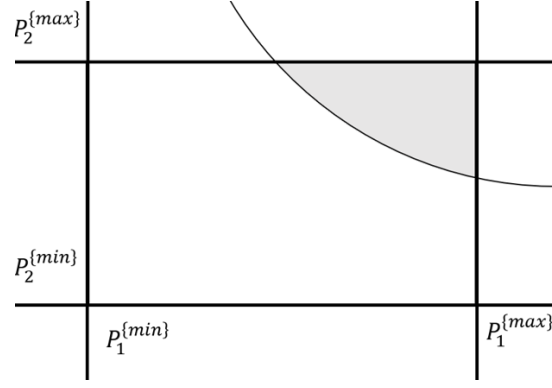


Figure 4.5.3: Schema of the relaxed convex set.

4.5.3 Feasible set: Linear relaxation

This linear approximation can also be used to replace the quadratic equality by linear constraints. Indeed, it is possible to use the information of the linear approximation to get a lower bound of this demand. Computing an accurate linear relaxation of this problem is very interesting since it provides a considerable reduction of the computational time.

By using the same normal vector as for the upper bound, a lower bound can be obtained by modifying the independent coefficient k such that all points on the demand curve are superior than this hyperplane. The representation of the feasible relaxed set can be illustrated on Figure 4.5.4. This is again a visual scheme for a simple problem with 2 units. The blue lines correspond to the linear upper and lower bound and the grey region consists in the feasible set. It is expected that the Pareto optimal values will lie on the lower bound constraint for the same reasons as in the latter section.

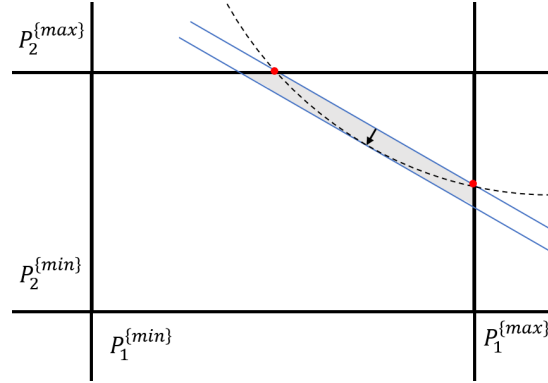


Figure 4.5.4: Linear relaxation of the domain.

In order to construct this linear relaxation, the first step is to compute the upper bound defined in equation (4.5.2). Since the lower bound is parallel to the upper bound, they therefore share the same normal vector $\mathbf{n}$. In addition, it is known that the gradient at each point $\mathbf{p}$ of the ellipsoid, defined as $\mathcal{E}$, is given by:

$$\nabla \mathcal{E}(\mathbf{p}) = 2B\mathbf{p} - \mathbf{e}.$$

Since the lower bound is defined as a hyperplane of normal vector $\mathbf{n}$ and intercepting only one point of the ellipsoid, the problem consists in finding $\mathbf{P}_0 \in \mathcal{E}$ such that $\nabla \mathcal{E}(\mathbf{P}_0) \parallel \mathbf{n}$. It can be

written as:

$$\mathbf{P}_0 \in [\mathbf{P}^{min}, \mathbf{P}^{max}]^N : \begin{cases} \mathbf{P}_0 = \frac{1}{2}B^{-1}(\alpha\mathbf{n} - \mathbf{e}) \\ \mathbf{P}_0^T B \mathbf{P}_0 - \mathbf{e}^T \mathbf{P}_0 + P_D = 0. \end{cases} \quad \alpha \in \mathbb{R} \quad (4.5.4)$$

Since this is complicated to solve analytically, point $\mathbf{P}_0$ can be obtained by solving the following optimization program:

$$\begin{aligned} \min_{\mathbf{p}} \quad & \sum_{i=1}^N n_i p_i \\ \text{s.t.} \quad & \mathbf{p}^T B \mathbf{p} - \mathbf{e}^T \mathbf{p} + P_D = 0. \\ & P_i^{min} \leq p_i \leq P_i^{max} \quad i = 1, \dots, N \end{aligned} \quad (4.5.5)$$

It corresponds to find the lowest hyperplane passing through the feasible domain. From this solution $\mathbf{P}_0$ we can now express the linear lower bound, with $k_{lower} = -\mathbf{n}^T \mathbf{P}_0$ as:

$$\sum_{i=1}^N n_i p_i + k_{lower} \geq 0. \quad (4.5.6)$$

4.6 Relaxation problems

The previous section explained some possibilities of relaxations but no formal proof has been established in order to show their validity for a general problem with feasible parameters. This section has therefore for objective to analyze the relaxations and to describe the exceptions that could arise. Three relaxations are considered, namely the convex relaxation, the linear relaxation using N points and a third relaxation is introduced that is a linear relaxation constructed from one feasible point.

4.6.1 Convex Relaxation

The first relaxation problem consists in relaxing the demand constraint from an equality to an inequality. The relaxed convex constraint is given by:

$$\mathbf{p}^T B \mathbf{p} - \mathbf{e}^T \mathbf{p} + P_D \leq 0 \quad (4.6.1)$$

where $\mathbf{e}$ represents the vector of ones. From section 4.4, it is shown that with the assumption that matrix B is positive definite, this equation represents the interior of an ellipsoid of degree N . Let us prove when the relaxation is exact.

The solution of the relaxation strongly depends on the gradient vector of the objective function. Therefore, the proof requires multiple preliminary observations before to be exposed. In the first time, an analysis of the behaviour of the gradient is performed. Then, 4 propositions are established on the problem and the introduced key points are illustrated on Figure 4.6.1 to have a graphic visualisation of these points. Proposition 4.6.5 concludes the section by demonstrating that the convex relaxation is exact given some conditions.

Let us first analyse when the objective is increasing monotonously. The scalarized objective function can be considered because the relaxed problem is convex. It has therefore the form $f(\mathbf{p}) = \lambda_1 E(\mathbf{p}) + \lambda_2 F(\mathbf{p})$. Let us analyze when $E(\mathbf{p})$ and $F(\mathbf{p})$ increase monotonously on the domain defined by the feasible box.

- The fuel costs are given by a quadratic function of type $F(\mathbf{p}) = \sum a_i + b_i p_i + c_i p_i^2$, and we have that $a_i, b_i, c_i \geq 0$. Therefore, all elements i of the gradient

$$\nabla[F(\mathbf{p})]_i = b_i + 2c_i p_i \quad (4.6.2)$$

are positive on the domain $[\mathbf{P}^{min}, \mathbf{P}^{max}]^N$. We have thus a monotone increase.

- The emission function is of type $E(\mathbf{p}) = \sum \alpha_i + \beta_i p_i + \gamma_i p_i^2 + \eta_i e^{\delta_i p_i}$. Here we have $\alpha_i, \gamma_i, \delta_i, \epsilon_i \geq 0$ but $\beta_i \in \mathbb{R}$. The i^{th} component of the gradient is given by

$$\nabla[E(\mathbf{p})]_i = \beta_i + 2\gamma_i p_i + \delta_i \eta_i e^{\delta_i p_i}. \quad (4.6.3)$$

We shall verify that it is positive on the domain, or in other terms that $\beta_i + 2\gamma_i P_i^{min} + \eta_i \delta_i e^{\delta_i P_i^{min}} \geq 0$ for each i . Unfortunately, this condition is not respected for the given parameters.

Let us define the unique output $\mathbf{P}_E$ that minimizes the unconstrained emission function, or such that $\nabla E(\mathbf{p}) = 0$. It corresponds to the solution of the equation:

$$\beta_i + 2\gamma_i P_{E,i} + \delta_i \eta_i e^{\delta_i P_{E,i}} = 0 \quad i = 1, \dots, N. \quad (4.6.4)$$

We do not have $\nabla E(\mathbf{p}) \succeq 0$ if at least one element i of the output satisfies $p_i < P_{E,i}$.

By considering the feasible box, the first requirement is therefore that $\nabla E(\mathbf{P}^{max}) \succeq 0$ in order to have $P_{E,i} \leq P_i^{max}$.

It is now possible to define the minimal power production set $\underline{\mathbf{P}}$ that respects the feasible box given by $[\mathbf{P}^{min}, \mathbf{P}^{max}]^N$. Its elements are specified by $\underline{P}_i = \max[P_i^{min}, P_{E,i}]$.

Proposition 4.6.1. Given a problem with N units and with fuel cost parameters $a_i, b_i, c_i \geq 0$, emission parameters $\alpha_i, \gamma_i, \delta_i, \epsilon_i \geq 0$, $\beta_i \in \mathbb{R}$ and $\nabla E(\mathbf{P}^{max}) \succeq 0$. The scalarized problem with $\lambda_1, \lambda_2 \geq 0$ defined by:

$$\begin{aligned} \min f(\mathbf{p}) &= \lambda_1 E(\mathbf{p}) + \lambda_2 F(\mathbf{p}) \\ \text{s.t. } \mathbf{p} &\in [\mathbf{P}^{min}, \mathbf{P}^{max}]^N \end{aligned} \quad (4.6.5)$$

has an optimal solution $\mathbf{P}^*$ that respects the condition that $\mathbf{P}^* \in [\mathbf{P}^{min}, \underline{\mathbf{P}}]^N$.

Proof. The optimal solution $\mathbf{P}^*$ has to be feasible, therefore: $P_i^* \geq P_i^{min}$ for each i .

We have to show that $P_i^* \leq \underline{P}_i$ for each i .

Remind that $\underline{\mathbf{P}}$ minimizes problem (4.6.5) with $\lambda_2 = 0$ and $\lambda_1 > 0$.

Given $\mathbf{P}_E$ the minimum of the unconstrained emission function: $\mathbf{P}_E = \arg \min E(\mathbf{p})$, let us define $\mathbf{P}_{EE}$ the minimum of the unconstrained scalarized objective: $\mathbf{P}_{EE} = \arg \min f(\mathbf{p})$.

If we can show that $P_{EE,i} \leq P_{E,i}$ for each i , then we are guaranteed that $P_i^* \leq \underline{P}_i$. Indeed, since $f(\mathbf{p})$ is increasing on each element i where $p_i \geq P_{EE,i}$, we have $P_i^* = \max[P_i^{min}, P_{EE,i}] \leq \underline{P}_i$.

Let us compute $P_{EE,i}$:

$$\begin{aligned} \nabla[f(\mathbf{p})]_i &= 0 \\ \Leftrightarrow \lambda_1(\beta_i + 2\gamma_i P_{EE,i} + \delta_i \eta_i e^{\delta_i P_{EE,i}}) + \lambda_2(b_i + 2c_i P_{EE,i}) &= 0 \\ \Leftrightarrow P_{EE,i} &= \frac{-\left(\lambda_1(\beta_i + \delta_i \eta_i e^{\delta_i P_{EE,i}}) + \lambda_2 b_i\right)}{2(\lambda_1 \gamma_i + \lambda_2 c_i)}. \end{aligned}$$

Consider that $P_{EE,i}$ is calculated and let us define $\zeta_i = \delta_i \eta_i e^{\delta_i P_{EE,i}}$.

If $\lambda_1(\beta_i + \zeta_i) + \lambda_2 b_i \geq 0$, then we are insured that $P_{EE,i} \leq 0$ and thus $P_i^* = P_i^{min}$. In this case element i respects the proposition.

On the other side: $\lambda_1(\beta_i + \zeta_i) + \lambda_2 b_i < 0$. Observe now that $(\beta_i + \zeta_i) < 0$.

Let us find out for which pair (λ_1, λ_2) we have maximal solution $P_{EE,i}$.

Define $\kappa = \frac{\lambda_1}{\lambda_2}$ with $\kappa \geq 0$. Then:

$$\begin{aligned} P_{EE,i} &= \frac{-(\kappa(\beta_i + \zeta_i) + b_i)}{2(\kappa\gamma_i + c_i)} \\ &= -\frac{\beta_i + \zeta_i}{2\gamma_i} + \frac{\overbrace{c_i(\beta_i + \zeta_i)}^{<0}}{2(\kappa\gamma_i + c_i)} - b_i \end{aligned}$$

is maximal for $\kappa \rightarrow \infty$, or when $\lambda_2 = 0$. Thus $P_{EE,i} \leq P_{E,i}$ and thus $P_i^* \leq \underline{P}_i$. $\square$

Remark. One can note on Figure 4.6.1 an illustration of a problem of 2 units with the displayed points $\underline{\mathbf{P}}$ and $\mathbf{P}^*$.

Consider problem (4.6.5). From Prop. 4.6.1, we know that $\mathbf{P}^* \in [\mathbf{P}^{min}, \underline{\mathbf{P}}]^N$. It is also guaranteed that $\nabla f(\mathbf{P}^*) \succeq 0$ because if an element of the gradient is negative, it is always possible to equal it to zero because each element of the gradient is independent to each other and because $\nabla[f(\mathbf{P}^{max})]_i \geq 0$ given that $\nabla[E(\mathbf{P}^{max})]_i \geq 0$.

Proposition 4.6.2. Consider symmetric positive definite loss matrix B in the constraint:

$$\mathbf{p}^T B \mathbf{p} - \mathbf{e}^T \mathbf{p} + P_D \leq 0. \quad (4.6.6)$$

Given $\mathbf{B}_i$ the i^{th} row of matrix B , if

$$2\mathbf{B}_i \mathbf{P}^{max} \leq 1 \quad i = 1, \dots, N. \quad (4.6.7)$$

then the demand function defined by $D(\mathbf{p}) = \mathbf{e}^T \mathbf{p} - \mathbf{p}^T B \mathbf{p}$ is increasing on the feasible box $[\mathbf{P}^{min}, \mathbf{P}^{max}]^N$.

Proof. Since $\mathbf{p} \in [\mathbf{P}^{min}, \mathbf{P}^{max}]^N$ and from condition (4.6.7), we can obtain:

$$\nabla[D(\mathbf{p})]_i = 1 - 2\mathbf{B}_i \mathbf{p} \geq 1 - 2\mathbf{B}_i \mathbf{P}^{max} \geq 0.$$

Therefore function D is increasing on the domain for every direction. Therefore, if $\mathbf{p}_1 \preceq \mathbf{p}_2$ then $D(\mathbf{p}_1) \leq D(\mathbf{p}_2)$. $\square$

Proposition 4.6.3. Consider the assumptions of Prop 4.6.1 and Prop. 4.6.2. Consider problem:

$$\begin{aligned} \min f(\mathbf{p}) &= \lambda_1 E(\mathbf{p}) + \lambda_2 F(\mathbf{p}) \\ \text{s.t. } \mathbf{p} &\in [\mathbf{P}^{min}, \mathbf{P}^{max}]^N \\ \mathbf{p}^T B \mathbf{p} - \mathbf{e}^T \mathbf{p} + P_D &\leq 0. \end{aligned} \quad (4.6.8)$$

If demand $P_D > \underline{D}$ with lower limit demand $\underline{D}$ defined by :

$$\underline{D} = \sum_{i=1}^N \underline{P}_i - \underline{\mathbf{P}}^T B \underline{\mathbf{P}} \quad (4.6.9)$$

then $\mathbf{P}^*$ is not feasible.

Proof. Recall that $\mathbf{P}^*$ is the solution of (4.6.5) and from Prop. 4.6.1, we have $\mathbf{P}^* \preceq \underline{\mathbf{P}}$. Recall from Prop. 4.6.2 the increasing demand function $D(\mathbf{p})$. Equation (4.6.6) can be written as:

$$D(\mathbf{p}) - P_D \geq 0.$$

Since $P_D > \underline{D} = D(\underline{\mathbf{P}})$ and $D(\mathbf{p})$ is increasing, we have:

$$D(\mathbf{P}^*) - P_D \leq D(\underline{\mathbf{P}}) - P_D < 0$$

and $\mathbf{P}^*$ is thus not feasible. $\square$

Remark. Note that the case with $P_D = \underline{D}$ is also valid. It remains true that $\mathbf{P}^* \preceq \underline{\mathbf{P}}$.

In the case $\mathbf{P}^* = \underline{\mathbf{P}}$, the relaxation is exact because: $(\underline{\mathbf{P}})^T B \underline{\mathbf{P}} - \mathbf{e}^T \underline{\mathbf{P}} + \underline{D} = 0$.

On the other hand $\mathbf{P}^* \neq \underline{\mathbf{P}}$, then either $D(\mathbf{P}^*) < \underline{D}$ and $\mathbf{P}^*$ is not feasible, or $D(\mathbf{P}^*) = \underline{D}$ and the relaxation is exact.

Therefore, we can use $P_D = \underline{D}$ in all cases. Either Prop. 4.6.3 is respected and it will be proved later that the relaxation is exact, or it directly insures that the demand constraint is tighten.

Proposition 4.6.4. By considering the solution $\mathbf{P}^{**}$ of problem (4.6.8). With the assumptions of Prop. 4.6.1 to 4.6.3, we are guaranteed that $\mathbf{P}^* \preceq \mathbf{P}^{**}$.

Proof. By contradiction, assume that $\exists i : P_i^{**} < P_i^*$.

It means therefore that $P_i^* > P_i^{min}$ and thus $P_i^* = P_{EE,i}$ since $P_i^* = \max[P_i^{min}, P_{EE,i}]$. Moreover, we have that $P_{EE,i} \leq P_i^{max}$ given the result from Prop 4.6.1 that $P_{EE,i} \leq P_{E,i} \leq P_i^{max}$ with assumption that $E(\mathbf{P}^{max}) \succeq 0$. Therefore $\nabla[f(\mathbf{P}^*)]_i = 0$.

Recall that $P_i^{**} < P_i^* = P_{EE,i}$ and let us therefore construct a better solution $\mathbf{P}^{better}$ with:

$$\begin{cases} P_j^{better} = P_j^{**} & \text{with } j \neq i \\ P_i^{better} = P_{EE,i} \end{cases}$$

that provides a lower objective value. This solution remains feasible in problem (4.6.8): it is in the feasible box since $P_i^{min} \leq P_{EE,i} \leq P_i^{max}$ and constraint (4.6.6) is respected given Prop. 4.6.2:

$$\mathbf{P}^{better} \succeq \mathbf{P}^{**} : D(\mathbf{P}^{better}) - P_D \geq D(\mathbf{P}^{**}) - P_D \geq 0.$$

Therefore, $\mathbf{P}^{better}$ is feasible for problem (4.6.8) and provides a lower objective value than $\mathbf{P}^{**}$ and the assumption is contradicted. $\square$

Remark. One can note on Figure 4.6.1 the representation of solution $\mathbf{P}^{**}$ of problem (4.6.8).

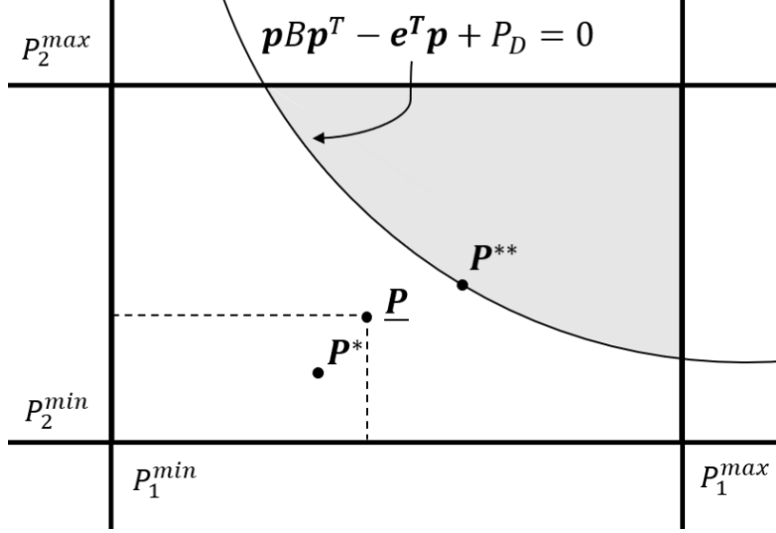


Figure 4.6.1: Schema of solutions $\underline{P}$, P^* and P^{**} for a problem with 2 units.

Proposition 4.6.5. Given the assumptions of Propositions 4.6.1 to 4.6.4, the solution of the unrelaxed problem:

$$\begin{aligned} \min f(\mathbf{p}) &= \lambda_1 E(\mathbf{p}) + \lambda_2 F(\mathbf{p}) \\ \text{s.t. } \mathbf{p} &\in [\mathbf{P}^{\min}, \mathbf{P}^{\max}]^N \\ \mathbf{p}^T B \mathbf{p} - \mathbf{e}^T \mathbf{p} + P_D &= 0. \end{aligned} \quad (4.6.10)$$

is equivalent to the solution of the convex relaxation problem (4.6.8).

Proof. Let us prove this by contradiction. Let us consider that the solution of problem (4.6.8) is not feasible for problem (4.6.10).

Suppose that P^{**} is the optimal solution of problem (4.6.8). Given the assumption, we therefore have:

$$(\mathbf{P}^{**})^T B \mathbf{P}^{**} - \mathbf{e}^T \mathbf{P}^{**} + P_D < 0. \quad (4.6.11)$$

P^{**} is optimal thus $\forall \mathbf{p} \in \mathcal{D} : f(\mathbf{P}^{**}) \leq f(\mathbf{p})$.

Let us construct a solution $\mathbf{P}^{\text{better}} = \mathbf{P}^{**} - d\mathbf{P}^i$ such that:

$$d\mathbf{P}^i \in \mathbb{R}^N := \begin{cases} \epsilon > 0 & \text{element } i \\ 0 & \text{else} \end{cases}$$

in a direction i where we have $P_i^{**} > P_i^*$. Recall from Prop. 4.6.4 that $P_i^{**} \geq P_i^*$, and suppose that $\mathbf{P}^{**} \neq \mathbf{P}^*$. The case $\mathbf{P}^{**} = \mathbf{P}^*$ is dealt later. $\mathbf{P}^{\text{better}}$ is thus constructed to provide a lower objective value than $\mathbf{P}^{**}$. Indeed, given that $\nabla f(\mathbf{P}^*) \succeq 0$, f is strictly monotonously increasing in the allowed directions i . If there exists at least one element i such that $\mathbf{P}^{\text{better}}$ is feasible, then the assumption that $\mathbf{P}^{**}$ is optimal is contradicted.

Since the domain is closed convex, there exists for each i a point $\mathbf{P}^{\text{better}}$ that satisfies constraint (4.6.6) because of assumption (4.6.11) but $\mathbf{P}^{\text{better}}$ is not guaranteed to be in the feasible box. Therefore, if $\mathbf{P}^{**} \neq \mathbf{P}^*$, we are guaranteed that for j such that $P_j^{**} > P_j^*$, there exists a better solution $\mathbf{P}^{\text{better}} = \mathbf{P}^{**} - d\mathbf{P}^j$ and $\mathbf{P}^{**}$ would not be optimal.

In the other case, $\mathbf{P}^{**} = \mathbf{P}^*$. Recall from the remark of Prop 4.6.3 that $P_D = \underline{D}$, otherwise $\mathbf{P}^*$ would not be feasible in the convex relaxation problem (4.6.8) and with the condition that $P_D \geq \underline{D}$.

Moreover, recall that $P_i^* \leq \underline{P}_i$ given Prop 4.6.1. Inequality (4.6.11) becomes :

$$\begin{aligned} (\mathbf{P}^*)^T B \mathbf{P}^* - \sum_{i=1}^N P_i^* + P_D &< 0 \\ \Leftrightarrow P_D &< D(\mathbf{P}^*) \leq D(\underline{\mathbf{P}}) = \underline{D} \end{aligned}$$

because the demand function $D(\mathbf{p})$ is increasing given Prop 4.6.2.

This contradicts the condition that $P_D = \underline{D}$ and therefore the initial assumption is contradicted. $\square$

Summary: The convex relaxation problem provides the exact solution of the unrelaxed problem if:

1. $\nabla[E(\mathbf{P}^{max})]_i \geq 0$, for $i = 1, \dots, N$.
2. $2\mathbf{B}_i \mathbf{P}^{max} \leq 1$, for $i = 1, \dots, N$.
3. $P_D \geq \underline{D}$ with $\underline{D}$ defined in (4.6.9).

4.6.2 Linear Relaxation using N extreme points

We are now considering the case where we replace the quadratic demand by a linear upper and lower bound starting from N extreme points. The method for computing these bounds is described in Section 4.5.3. However, the method for obtaining N extreme points for a general problem is not considered and is described in this section. It is then followed by a critical analysis of the validity of this relaxation model.

N extreme points

Section 4.5.3 assumed that each edge of the feasible box passing through coordinate $\mathbf{P}^{max}$ has an intersection with the ellipsoid, known as the demand curve. However, this assumption is not respected for most of the demand amounts. We need therefore to describe and analyze the method for obtaining the N extreme points.

In order to approximate the upper bound, we are looking for N points located on the extremity of the feasible box. It is considered that there exists at least one edge in each direction p_i such that the method can be looped on every direction. For each direction k , an edge intercepting the ellipsoid can be defined as a sequence $\mathbf{S}(z) = [S_1, S_2, \dots, S_k = z, \dots, S_N]$ where

$$S_i = \begin{cases} P_i^{min} \\ P_i^{max} \end{cases} \quad i = 1, \dots, N, \quad i \neq k. \quad (4.6.12)$$

The condition that this edge, or sequence, crosses the demand function inside the feasible box is given by:

$$\exists z \in [P_k^{min}, P_k^{max}] : \mathbf{S}(z)^T B \mathbf{S}(z) - \mathbf{e}^T \mathbf{S}(z) + P_D = 0 \quad (4.6.13)$$

This sequence can be found by solving the following binary optimization problem:

$$\begin{aligned}
 & \min_{\mathbf{x}, \mathbf{S}} 0 & (4.6.14) \\
 & \text{s.t. } S_i = x_i P_i^{max} + (1 - x_i) P_i^{min} & i = 1, \dots, N, i \neq k \\
 & \mathbf{S}(P_k^{max})^T B \mathbf{S}(P_k^{max}) - \mathbf{e}^T \mathbf{S}(P_k^{max}) + P_D \leq 0 \\
 & \mathbf{S}(P_k^{min})^T B \mathbf{S}(P_k^{min}) - \mathbf{e}^T \mathbf{S}(P_k^{min}) + P_D \geq 0 \\
 & x_i \in \{0, 1\} & i = 1, \dots, N
 \end{aligned}$$

The algorithm is represented graphically on Figure 4.6.2 for a problem with 3 generators. At each iteration we have to find a sequence, or edge, in the direction k that crosses the ellipsoid. The orange lines represent all edges of direction k , and the red points correspond to the intersection points.

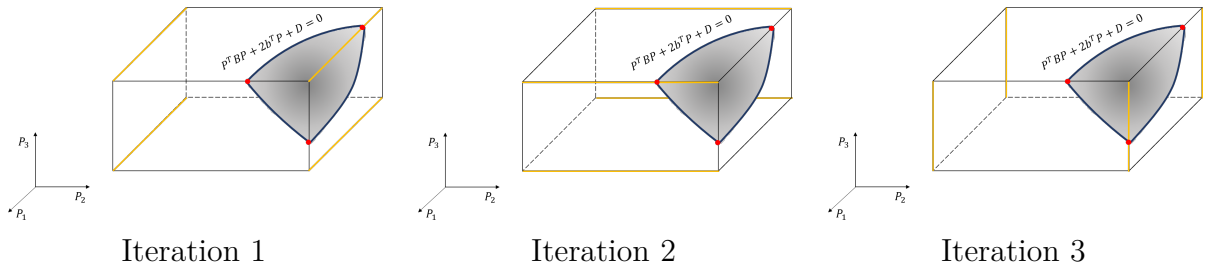


Figure 4.6.2: Search of extreme points following each axis for a 3-unit problem. The grey region represents the demand function, the box is defined as $[\mathbf{P}^{min}, \mathbf{P}^{max}]^3$, the orange lines are the candidate sequences at each iteration.

This algorithm is not robust for each form of the demand. It has indeed been assumed that the feasible box has at least one intersection with its edge for each direction. In other words, it is supposed that each generator is able to produce at minimal or maximal rate with the other generators arranging their production to respond to the demand. An geometric illustration is displayed on Figure 4.6.3.

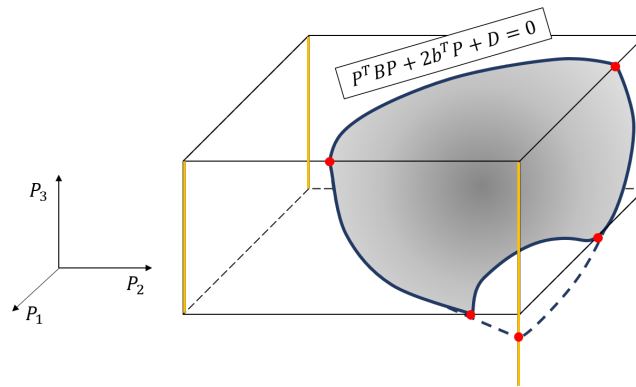


Figure 4.6.3: Case where the search method fails.

Let us analyze the condition to prevent such problems considering direction k . Note that this analysis has to be verified for all edge directions $k = 1, \dots, N$. Firstly, we have that there exists 2^{N-1} sequences/edges in direction of k . Call $\mathbf{S}^j(z_k)$ the j^{th} sequence in direction k and defined as:

$$\mathbf{S}^j(z_k) = [S_1^j, S_2^j, \dots, S_k^j = z_k, \dots, S_N^j].$$

We can define the demand that sequence j can fulfill with $P_k^{min} \leq z_k \leq P_k^{max}$:

$$D_j(z_k) = D[\mathbf{S}^j(z_k)] = \sum_{i=1}^N S_i^j - \mathbf{S}^j(z_k)^T B \mathbf{S}^j(z_k).$$

If we classify $D_j(z_k)$ in increasing order by its lowest production generation, i.e., when $z_k = P_k^{min}$, the condition is that:

$$\exists l \leq j : D_l(P_k^{max}) \geq D_{j+1}(P_k^{min}) \quad j = 1, \dots, 2^{N-1}.$$

On Figure 4.6.4 is illustrated as an example a representation of the range of demands that can be covered by the sequences. We note that the condition is not respected before D_{m+1} because the red zone is not covered by D_m . We therefore have that for demands that are between $D_m(P_k^{max})$ and $D_{m+1}(P_k^{min})$, the ellipsoid is such that it does not intercept any edge of direction k in the feasible box $[\mathbf{P}^{min}, \mathbf{P}^{max}]^N$.

However, we can still use the point that is on the prolongation of the edge described by $\mathbf{S}^m(z_k)$ for having a quite accurate upper bound.

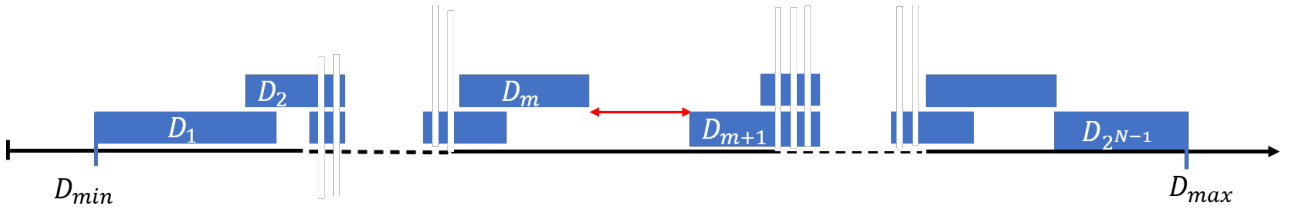


Figure 4.6.4: Graphic representation on the demand that all sequences can cover for a direction k .

Once the N points are calculated, the linear upper bound (4.5.2) and lower bound (4.5.6) can be constructed as described in section 4.5.3.

Drawbacks of this method

We note that this method requires to solve N integer programs because each extreme point is found by solving a particular optimization problem. In addition, another non-convex minimization problem has to be solved for calculating the lower bound. The first drawback is therefore the computational time. Since the goal of relaxations is to make the problem easier to solve, the run time for obtaining and solving the linear relaxation has to be taken into account. We will indeed compare the resolution times for all relaxations in Section 4.7.2.

Moreover, this technique has also to be criticized in the sense that the N extreme points are chosen randomly from all possible extreme points. We know that there are 2^{N-1} candidate edges in each direction, but in many cases, there is no unique solution to problem (4.6.14). It depends on the total demand as it is shown on the previous figure. With other words, for a given demand, how much sequences could respond to this demand? The following table presents results for some experiences with the 4 smallest test cases:

N	Demand [MW]	Maximal production [MW]	Intersections	Total edges
2	100	343.15	2	4
	200		2	
	300		2	
3	140	490.08	3	12
	280		6	
	420		3	
6	400	1290.75	6	192
	800		64	
	1200		6	
10	666.66	2259.4	15	5120
	1333.33		1061	
	2000		189	

Table 4.1: Analysis of the total number of intersections with different test cases

It is difficult to establish a simple correlation between the number of intersections, the problem size and the demand but some observations can be noted.

For a sufficiently high demand requiring $N - 1$ generators producing at maximal generation, or inversely, for a low demand requiring the minimal generation for $N - 1$ generators, there are exactly N intersections between the feasibility box and the demand function. This can be illustrated for the 2, 3 and 6-unit test case with low or high demand. Another note is that as the problem size increases, the total number of intersections increases non linearly.

As a conclusion, the choice of the extreme points matters in a way that it is not guaranteed that their linear combination is superior to all feasible points. Indeed, if there are more than N intersections, there exists only one combination of N extreme points that provides a linear upper bound for all feasible points. In this case, the feasible set is known as the convex hull of the domain. However, if another combination of N points is chosen, there are still chances that it bounds superiorly the optimal solution of the unrelaxed problem and making the relaxation acceptable. It has therefore to be considered as a heuristic.

On the other hand, as we saw in Section 4.6.1, the Pareto optimal values are often constrained by the lower bound and the upper bound is not required in most of the cases. The use of the lower bound is therefore sufficient and the valid linear relaxed set is illustrated on Figure 4.6.5 corresponding to the grey region for a 2 dimensional problem.

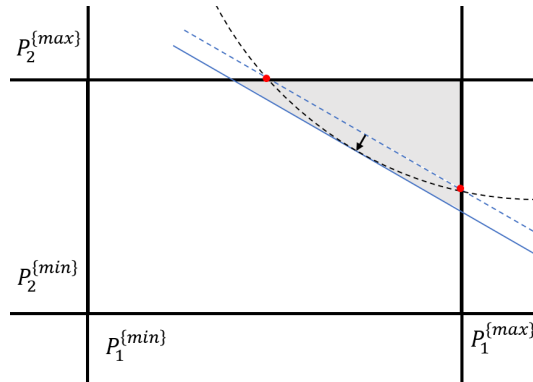


Figure 4.6.5: Relaxed linear set valid for all problems.

4.6.3 Linear relaxation from 1 feasible point

From the previous section, it has been discovered that it is not very efficient to start from an upper bound to create a lower bound. This section describes another method for obtaining a linear relaxation of the set and is called the linear relaxation from 1 feasible point. This relaxation is introduced by Loïc Van Hoorebeeck (LVH) in [26].

The method consists in computing in the first time a lower bound. The ellipsoidal curve in the feasible box is reasonably flat because the ratios between the size of the feasible box and the radii of the ellipsoid of degree N are very small (less than 1.5% for the 10-unit test case). Therefore, the normal vector of the feasible points at the ellipsoid does not vary much. The normal vector $\mathbf{n}$ of the ellipsoid is given by:

$$\mathbf{n} = -\nabla \mathcal{E}(\mathbf{P}_0) = -(2B\mathbf{P}_0 - \mathbf{e}). \quad (4.6.15)$$

Therefore, vector $\mathbf{n}$ varies at most of:

$$\Delta \mathbf{n} = 2B(\mathbf{P}^{max} - \mathbf{P}^{min}) \ll \mathbf{e}.$$

The method consists in taking a random feasible point $\mathbf{P}_0$ and obtaining a lower bound with the normal at this point on the ellipsoid. Note that the opposite direction is taken in order to have a vector that points inside the ellipsoid. This normal vector is also used for the upper bound. As in the previous section, the upper and the lower bound constraint differs only in their independent coefficient k . We have:

$$k_{lower} = -\mathbf{n}^T \mathbf{P}_0 \quad (4.6.16)$$

for the lower bound. Coefficient k_{upper} is computed such that every feasible point of the quadratic demand is lower than the linear upper bound. This coefficient can be found by modifying problem (4.5.5) into a maximization problem. The linear relaxation of the set is illustrated on Figure 4.6.6 and the grey area represents the relaxed feasible set.

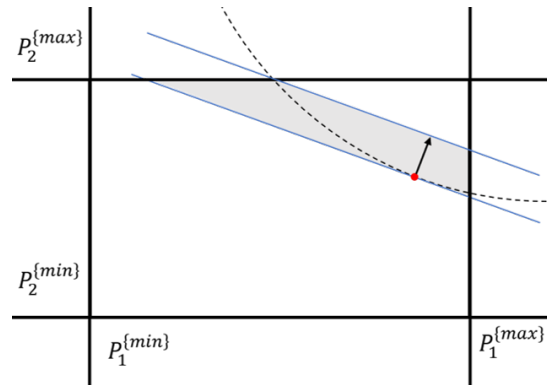


Figure 4.6.6: Linear relaxation from 1 feasible point.

4.7 Comparison of the relaxations

Section 4.6 described three methods for relaxing the non-convex problem. It has been shown that the convex relaxation consisting in relaxing the equality into an inequality can be exact for well chosen parameters. On the other side, the two linear relaxations provide lower bounds on the objective values. This section aims at analyzing these gaps while also comparing the computational times.

4.7.1 Error analysis

As already described in Section 1.6.1, the error analysis for a multi-objective optimization problem with objectives of different units is hard to express. Three methods are explained for computing these errors for the linear relaxations and are applied on the 6-unit test case for three different demands.

The first technique is already explained in section 1.6.1 for analyzing the Hermite interpolation. It consists in computing the distance following the two axes defined by the objectives. Let us define the name of this method as the error following the axes. From Figure 4.7.1, the errors are defined by E_{f_1} and E_{f_2} for a given point. They represent the error made when we fix one objective value to the same constant and compare the difference of the second objective.

Secondly, the next method is called error with minimal distance. In order to make both functions comparable by working on the same order of magnitude, it can be considered that one pound is equal to π dollar, defined by equation 1.3.4. From this adjustment, the nearest point between a given point of the exact front and the relaxed Pareto curve can be computed in terms of Euclidean norm. A similar approach can be found in the literature by normalizing the objective functions [27]. Recall that the distance between two objectives associated to the Pareto values $\mathbf{p}_1$ and $\mathbf{p}_2$ in the Euclidean norm is defined by:

$$d(\mathbf{p}_1, \mathbf{p}_2) = \left([\pi E(\mathbf{p}_1) - \pi E(\mathbf{p}_2)]^2 + [F(\mathbf{p}_1) - F(\mathbf{p}_2)]^2 \right)^{\frac{1}{2}}. \quad (4.7.1)$$

Once again, these two points are compared following their emission and cost components. On Figure 4.7.1, the errors are defined as ϵ_{f_1} and ϵ_{f_2} for a given point.

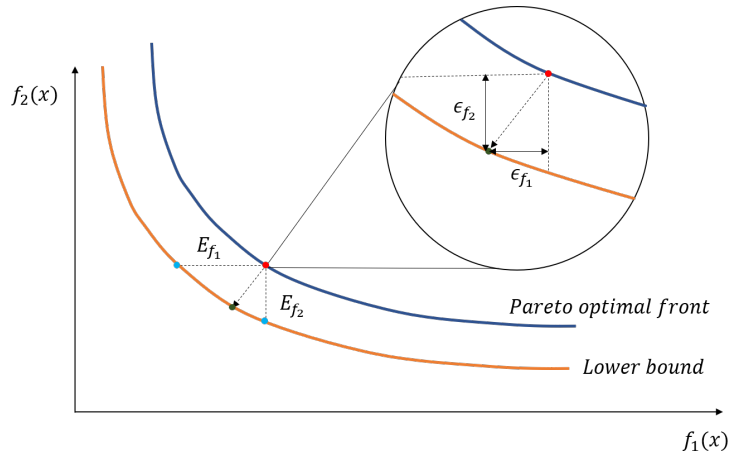


Figure 4.7.1: Two methods for computing the error at a given point (red dot).

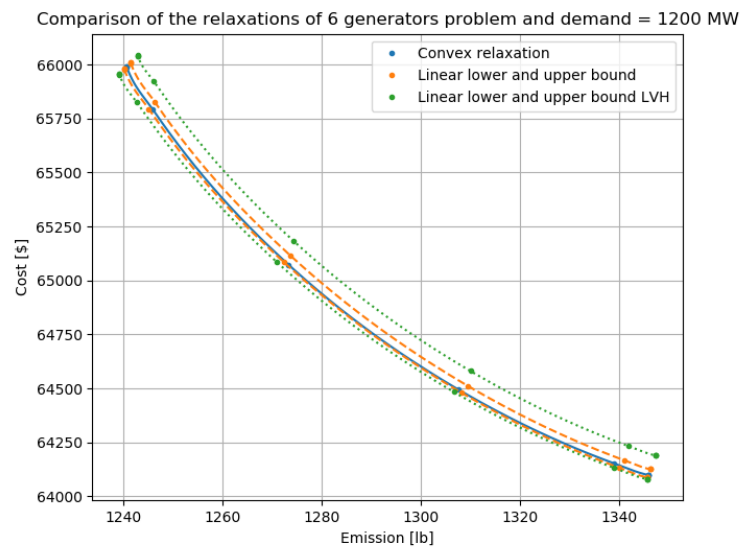
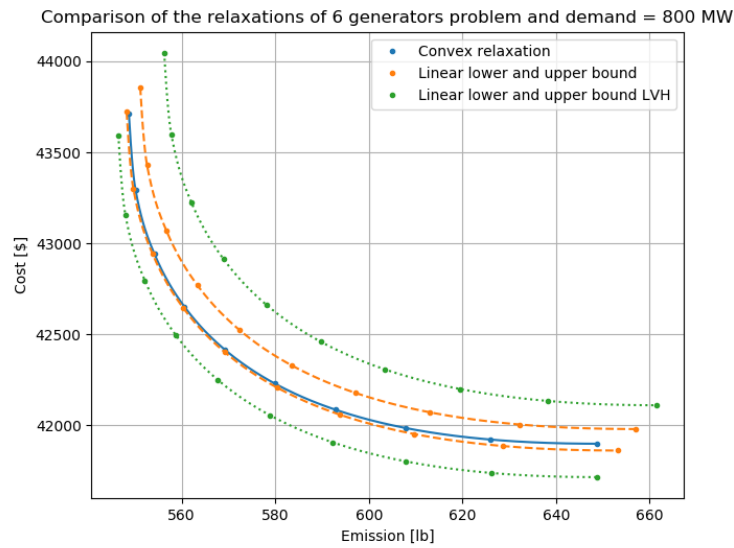
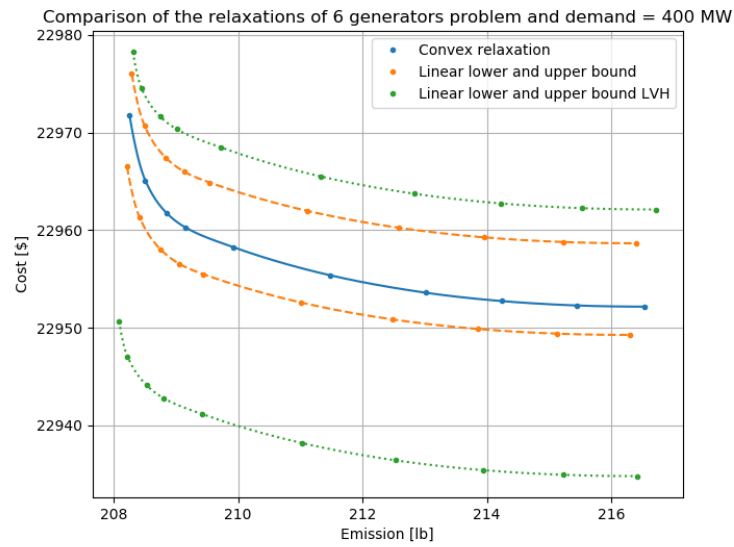


Figure 4.7.2: Pareto optimal front of the relaxations for the 6-unit problem.

The last method is computed by comparing the Pareto optimal values for the same scalarized objective function. For sake of simplicity, this error is defined as the error on the w -objective. In order to obtain the error following each objective, we compute the difference between each objective value at the Pareto optimal point of the exact and the relaxed problem.

Some observations can be made from these three methods for different examples. The 6-unit test case is used to illustrate them for 3 different demands (400, 800 and 1200 MW) and are represented on Figure 4.7.2. Note at first that the convex relaxation does not present any errors since it is equivalent to the non-convex problem for the used parameters. This relaxation is therefore not analysed in terms of error.

Two Pareto fronts are displayed in the same color corresponding to the feasible set above the lower bound and the overlying front is obtained with the feasible set above the upper bound. Unfortunately, the POF's for the relaxations provide sometimes a large gap, as it can be noted on the first graph of the latter figure.

From this graph, the first method for measuring the error is not representative because it compare points that doesn't have the same signification. Indeed, for example, the solutions close to the Economic Dispatch solution of the unrelaxed problem are compared to the Environmental Dispatch solution of the of linear relaxation using one feasible point. Hence, the second and third method are more pertinent. Since relaxation problems focus more on each scalarized problem independently and less on the Pareto front, the analysis for the relaxations is more suited by using the error on the w -objective.

The errors for the 3 examples are illustrated on each of the three diagrams of Figure 4.7.3. Each diagram represents the combined maximum relative error of one method. In order to present the error on both the emissions and the costs, they are displayed as the mean value between these two relative errors. Therefore, it is noticed from the three methods that the two relaxations generates more relative error on the emissions than on the costs. From all these diagrams, it can be easily noticed that the relaxation using N points generates less error than the relaxation using one feasible point. It is definitely shown that the POF of this relaxation is closer than the one using 1 feasible point via the error on the minimal distance. By considering the error on the w -objective, the relative error of the relaxation using N points varies between 0.07% and 0.69% for the emissions. The error on the costs is less important relatively since it fluctuates between 0.02% and 0.09%. On the other side, the relaxation constructed from 1 feasible point generates a relative error of 0.24-0.4% on the emissions and 0.05-0.44% error on the costs.

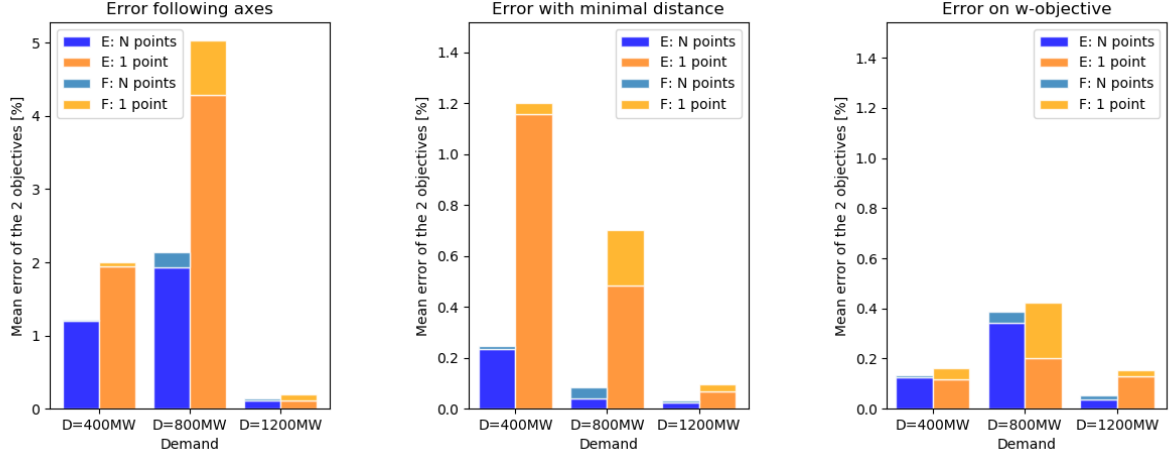


Figure 4.7.3: Combined relative error of the emissions and the costs for the two relaxations using different methods.

4.7.2 Computational time

The previous analysis described the error that the relaxations present but another important measure is its time efficiency. Remind that three relaxations need to be compared : the convex relaxation, the linear relaxation using N points and the relaxation based on one feasible point. On one side, the convex relaxation problem has for complexity having a quadratic inequality in its constraints. On the other side, the two linear relaxations need to solve the additional problem for obtaining the linear constraint. This section describes therefore the time efficiency of these relaxations.

First of all, the time for obtaining the coefficients of the linear constraints for the two relaxations has to be compared. The relaxation using N points requires to solve N mixed integer problems followed by a non-convex minimization program to get the lower bound. On the other hand, the relaxation using one point performs one non-convex feasibility problem for finding the initial point P_0 , followed by a non-convex maximization program to get the upper bound.

We will see in Section 4.9 that **Gurobi** does not manage very well nonlinear constrained optimization programs when the problem size increases. Another solver that is more suited is **Scipy**. It can indeed be used for obtaining the initial feasible point of the relaxation. As a result, Figure 4.7.4 shows the computational time for computing the linear coefficients for problems of different sizes. Firstly, it can be observed that the computational time grows exponentially for the two methods using **Gurobi**. The relaxation using 1 point is slightly more efficient but remains of the same order. Lastly, it can be noticed that the **Scipy** solver has much better performances when the size of the problem increases.

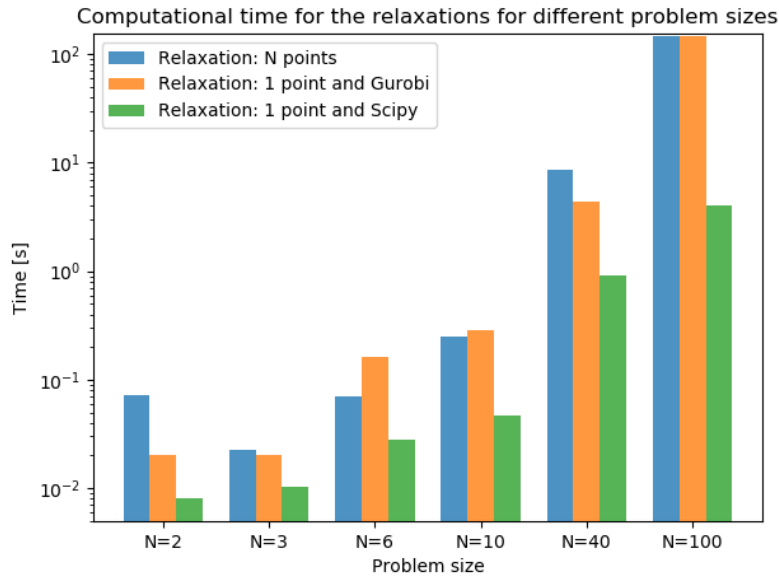


Figure 4.7.4: Computational time for obtaining the linear constraints.

Then, let us analyze the difference between the computational time of the convex and linear relaxation on **Gurobi** and the non-convex solver on **Scipy**. The results are shown on Figure 4.7.5 and present the run time for computing one single optimization problem. Remind that for obtaining an estimation of the Pareto front, several problems have to be considered. **Gurobi** is indeed very efficient for small linear constrained problems and its run time can even be neglected. On another side, it can also be confirmed that **Scipy** performs better than **Gurobi** for larger problems.

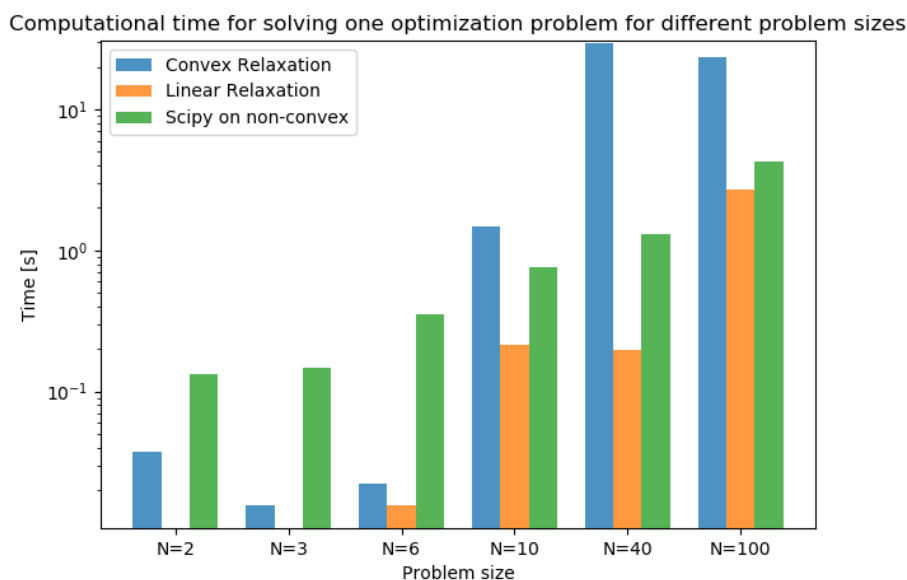


Figure 4.7.5: Computational time for solving one optimization problem.

As a conclusion, the linear relaxation using N points provides a smaller gap with the exact problem compared to the relaxation constructed from one feasible point. Even if the run time of the relaxation with N points is slightly slower, it remains quite equivalent with the relaxation with one point. For these reasons, the relaxation using N points present better performances than the latter.

However, this relaxation cannot compete with nonlinear solvers on `Scipy` for this type of problem. Indeed, the problem is relatively simple and contains only two types of constraints: namely the generation limits and the quadratic demand.

4.8 First order solvers

From Section 4.3, it has been shown that methods based on the gradient vector as search direction can be efficient, even on the non-convex problem. Moreover, it is shown in 4.6.1 that the convex relaxation of the EED using the transmission losses is exact given certain conditions on the parameters. Therefore, this section presents the results for 3 first-order iterative methods on the convex problem and analyzes their efficiency. The Gradient Method, the Accelerated Gradient and the Spectral Projected gradient are tackled in this section.

4.8.1 Gradient and Accelerated Method

Recall that an initial analysis is done in Section 4.3 on the Gradient Method and the Accelerated method on the non-convex problem. This analysis can be done on the convex relaxation but the convergence is now guaranteed since a convex set is considered. Therefore, this section computes and compares the convergence of these two models with their theoretical rate of convergence.

Firstly, recall that the Gradient Method and the Accelerated method are described in algorithm (4.3.2) and (4.3.3) respectively. The convergence of these two methods has to be compared with their theoretical rate in order to characterize their general efficiency. From [24], each iterate $\mathbf{P}_k$, obtained by applying the Gradient Method with optimal step size $h = \frac{2}{\mu+L}$, insures that:

$$f(\mathbf{P}_k) - f^* \leq \frac{L}{2} \left(\frac{L - \mu}{L + \mu} \right)^{2k} \|\mathbf{P}_0 - \mathbf{P}^*\|^2. \quad (4.8.1)$$

On the other side, the convergence of the Accelerated Method for a μ -strongly convex function is given by:

$$f(\mathbf{P}_k) - f^* \leq 2 \left(1 - \sqrt{\frac{\mu}{L}} \right)^k (f(\mathbf{P}_0) - f). \quad (4.8.2)$$

On Figure 4.8.1 is displayed the real and theoretical rate of convergence of the Gradient and Accelerated Method for the 10-unit test system. We note that for this specific problem, the Gradient Method presents a better theoretical rate. In order to compare these rates, let us have a look at the maximal number of iterations.

Given the condition number $\kappa = \frac{L}{\mu} > 1$, the inequality: $1 - \tau \leq e^{-\tau}$ can be used by considering $\tau = \frac{2}{1+\kappa}$. Assuming that the right-hand side of inequality (4.8.1) is smaller than ϵ , we get the following bound for the number of steps:

$$N \geq \frac{1}{4} (1 + \kappa) \ln \left(\frac{L \|\mathbf{P}_0 - \mathbf{P}^*\|^2}{2\epsilon} \right). \quad (4.8.3)$$

Using the same trick for the Accelerated method, we get

$$N \geq \sqrt{\kappa} \ln \left(\frac{2(f_0 - f^*)}{\epsilon} \right). \quad (4.8.4)$$

In general κ is big so the second bound is much better. However, in the 10-unit test case, parameter κ is relatively small leading to a slightly better bound of the Gradient Method.

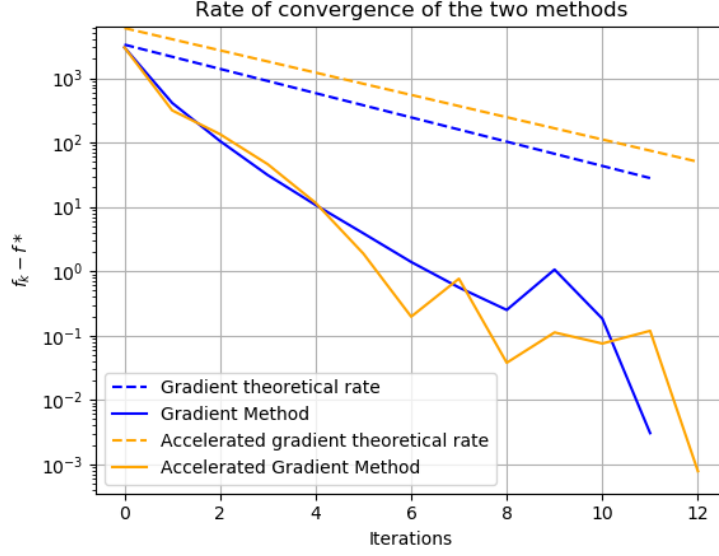


Figure 4.8.1: Convergence of the relaxed convex problem with 10 generators.

From this observation, the two methods present good performances since they require only 11 and 12 iterations for having an accuracy of less than 10^{-2} . Moreover, the run time of the convex relaxed solution obtained with these first-order methods are better than the direct implementation in *Gurobi*. Indeed, the Accelerated Method takes about 0.4 seconds while *Gurobi* takes approximately 1.5 seconds. However, despite the fact that they are efficient on this small problem, the solver encounters numerical issues while computing the projection for larger problems as the 40-unit and 100-unit test case.

4.8.2 Spectral Projected Gradient

The third method that is introduced is the Spectral Projected Gradient (SPG) algorithm. This method uses also the gradient vector as search direction like the two previous methods. However, it uses an optimized step length that is based on the strategy proposed by Barzilai and Borwein for the unconstrained minimization of a differentiable function. The SPG is designed for convex constrained optimization and the algorithm is given in [28] :

Algorithm 2 SPG

```

1:  $\lambda_{min} = 10^{-10}, \lambda_{max} = 10^{10}$ 
2:  $\lambda_0 = \frac{1}{L}, \mathbf{P}_0 \in \mathcal{D}$ 
3: for  $k = 1, \dots$  do
4:    $\mathbf{d}_k = \text{Proj}_{\mathcal{D}}[\mathbf{P}_k - \lambda_k \nabla(f_k)] - \mathbf{P}_k$ 
5:   Obtain  $\alpha_k$  from algorithm 3.
6:    $\mathbf{P}_{k+1} = \mathbf{P}_k + \alpha_k \mathbf{d}_k$ 
7:    $\mathbf{s}_k = \mathbf{P}_{k+1} - \mathbf{P}_k$ 
8:    $\mathbf{y}_k = \nabla(f_{k+1}) - \nabla(f_k)$ 
9:   if  $\mathbf{s}_k^T \mathbf{y}_k \leq 0$  then
10:     $\lambda_{k+1} = \lambda_{max}$ 
11:   else
12:     $\lambda_{k+1} = \max \left\{ \lambda_{min}, \min \left( \frac{\mathbf{s}_k^T \mathbf{s}_k}{\mathbf{s}_k^T \mathbf{y}_k}, \lambda_{max} \right) \right\}$ 
13:   end if
14: end for
    
```

Algorithm 3 Spectral step length α_k

```

1:  $\sigma_1 = 0.1, \sigma_2 = 0.9, \gamma = 10^{-4}, M = 10$ 
2: Given  $k, \alpha = 1$ 
3:  $f_{max} := \max_{k+1-M \leq i \leq k} f_i$ 
4:  $\tilde{\mathbf{P}} := \mathbf{P}_k + \alpha \mathbf{d}_k$ 
5: while  $(f(\tilde{\mathbf{P}}) > f_{max} + \gamma \alpha \nabla(f_k)^T \mathbf{d}_k)$  do
6:    $\bar{\alpha} = \frac{-\frac{1}{2} \alpha^2 \nabla(f_k)^T \mathbf{d}_k}{f(\tilde{\mathbf{P}}) - f_k - \alpha \nabla(f_k)^T \mathbf{d}_k}$ 
7:   if  $\bar{\alpha} \in [\sigma_1, \alpha \sigma_2]$  then
8:      $\alpha \leftarrow \bar{\alpha}$ 
9:   else
10:     $\alpha \leftarrow \frac{\alpha}{2}$ 
11:   end if
12:    $\tilde{\mathbf{P}} \leftarrow \mathbf{P}_k + \alpha \mathbf{d}_k$ 
13: end while
14:  $\alpha_k := \alpha$ 
    
```

The spectral length is sensitive to certain parameters and are considered here as the default values. They only require to satisfy the conditions that $\gamma \in (0, 1)$, $0 < \sigma_1 < \sigma_2 < 1$ and that M is integer with $M \geq 1$. The additional assumptions are on the safeguarding parameters $0 < \lambda_{min} \leq \lambda_{max} < \infty$ for the spectral step length. The initial spectral step length λ_0 is arbitrary with $\lambda_0 \in [\lambda_{min}, \lambda_{max}]$. Therefore, it is chosen that $\lambda_0 = \frac{1}{L}$ since it is an accurate step length given the Gradient step and it is fast to calculate. The convergence of this method is illustrated on Figure 4.8.2 for the 10-unit and 40-unit test case. The computational time for solving the convex relaxation with the SPG is approximately 0.5 seconds for the 10-unit system comparing to 0.4 seconds for the Accelerated Method and 1.5 seconds for **Gurobi**. The 40-unit problem takes 3.4 seconds comparing to 40 seconds for **Gurobi**.

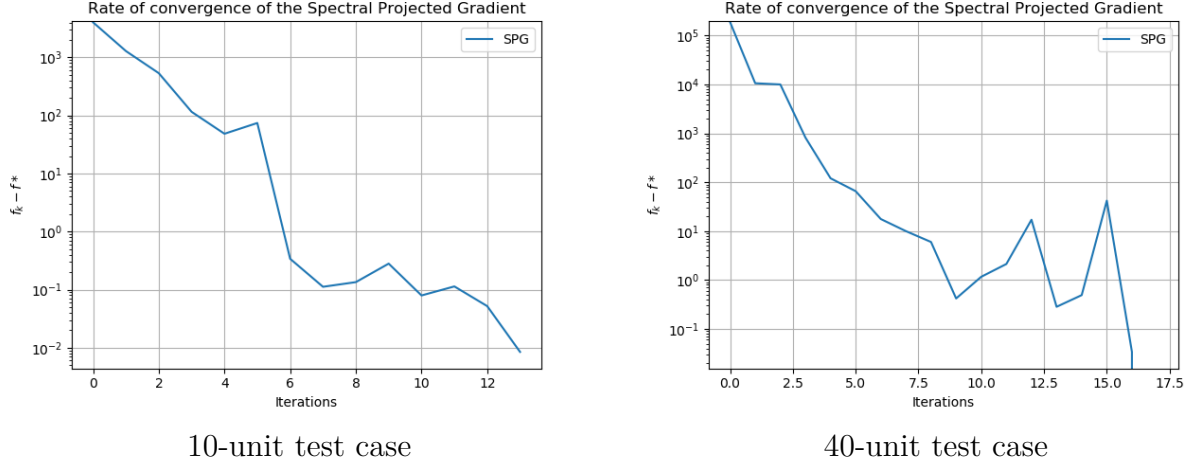


Figure 4.8.2: Convergence of the Spectral Projected Gradient.

In order to conclude this section, it is observed that the SPG presents the best performances for solving the problem. Indeed, comparing to the direct implementation of the problem in **Gurobi**, it is able to solve in a very efficient way the convex relaxation of the static EED. In addition, this first-order method can handle the 40-unit problem whereas the Gradient and Accelerated Method are not able to do so. However, it remains difficult to obtain the feasible projections of the iterates for large problem sizes even with the SPG. Indeed, a 100-unit test system is therefore created (see Appendix A.6) and it is observed that numerical troubles are encountered during this step. Therefore, these first-order methods present their weakness in the fact that they cannot deal with large-scale problems. Additionally, it will be shown in 4.9.1 that the time efficiency can be increased even more.

4.9 Second order methods

We notice that the scalarized objective function of the EED has extremely interesting properties. They can be exploited in order to optimize the solver and make it robust when the size increases. Second order methods are based on the second derivative of the functions to enhance the convergence. Since the Hessian matrix of the objective function is diagonal, it is very easy to obtain its inverse and this parameter can be used to improve the program.

Let us first notice that **Scipy** has already some nonlinear solvers implemented for nonlinear constrained optimization [29]. These methods are therefore able to manage non-convex programs and are interesting to cite here.

The first method is the Sequential least squares programming algorithm ("SLSQP"). It can manage easily the quadratic equality constraint and is much more efficient than **Gurobi**, even for the convex relaxed problem. This method can also treat bigger problems like the 40 or 100-unit problem in the static problem.

When the problem size increases, the Trust-Region Constrained Algorithm method ("trust-constr") is more efficient. It requires more information on the problem, namely the first and second derivatives of the constraints. For the larger problems with more than 10 generators, this method is faster than the latter.

Note that these two solvers in **Scipy** are efficient only on the static version of the problem. These solvers are not well suited for highly constrained problems because they rely on the initial guess of the solution. By considering the convex relaxation of the dynamic EED, **Gurobi** is able to solve one problem approximately 13 times faster than **Scipy**. This number is based on comparison for the 10-unit test case: it takes 22 seconds for **Gurobi** and 290 seconds for **Scipy** with the *trust-const* method.

4.9.1 Sequential Quadratic Programming

It is possible to decrease the computational time by implementing the general sequential quadratic programming (SQP) algorithm on the non-convex problem. By considering the problem written as:

$$\begin{aligned} \min_{\mathbf{p}} \quad & f(\mathbf{p}) \\ \text{s.t.} \quad & h(\mathbf{p}) = 0 \\ & g(\mathbf{p}) \leq 0 \end{aligned} \quad (4.9.1)$$

where f is the scalarized objective of the costs and the emissions, h is the quadratic constraint on the demand and $g : \mathbb{R}^N \rightarrow \mathbb{R}^{2N}$ represents the power generation restrictions. Given the Lagrangian multipliers λ and $\boldsymbol{\mu}$ for h and g respectively, the Lagrangian function is given by:

$$\mathcal{L}(\mathbf{p}, \lambda, \boldsymbol{\mu}) = f(\mathbf{p}) - \lambda h(\mathbf{p}) - \boldsymbol{\mu}^T g(\mathbf{p}). \quad (4.9.2)$$

This second order method computes an optimal step $\mathbf{d}$ at each iteration by solving the minimization problem given in [30] that corresponds to:

$$\begin{aligned} \min_{\mathbf{d}} \quad & \nabla f(\mathbf{P}^k)^T \mathbf{d} + \frac{1}{2} \mathbf{d}^T \nabla_{pp}^2 \mathcal{L}(\mathbf{P}^k, \lambda^k, \boldsymbol{\mu}^k) \mathbf{d} \\ \text{s.t.} \quad & h(\mathbf{P}^k) - \nabla h(\mathbf{P}^k)^T \mathbf{d} = 0 \\ & g(\mathbf{P}^k) - \nabla g(\mathbf{P}^k)^T \mathbf{d} \leq 0 \end{aligned} \quad (4.9.3)$$

However, this search direction problem is difficult to use since it requires to compute at each step the Lagrangian multipliers. It is first noticed that all inequality constraints g_i are linear, leading to a second derivative equal to zero. Then, it can be considered that the equality constraint h does not affect much the Hessian of the Lagrangian function and can be removed. By doing this, the approximation can be seen as a linear approximation of the demand constraint instead of a quadratic approximation, done initially by the SQP method. It will be observed later that this approximation converges to the solution in an efficient time and with less iterates. The search subproblem can therefore be written as:

$$\begin{aligned} \min_{\mathbf{d}} \quad & \nabla f(\mathbf{P}^k)^T \mathbf{d} + \frac{1}{2} \mathbf{d}^T \nabla^2 f(\mathbf{P}^k) \mathbf{d} \\ \text{s.t.} \quad & h(\mathbf{P}^k) - \nabla h(\mathbf{P}^k)^T \mathbf{d} = 0 \\ & P_i^{\min} \leq P_i^k + d_i \leq P_i^{\max} \quad i = 1, \dots, N \end{aligned} \quad (4.9.4)$$

The SQP algorithm goes as follows:

$$\mathbf{P}^0 \in \mathcal{D} \quad (4.9.5)$$

Obtain search direction $\mathbf{d}^k$ by solving problem (4.9.4)

$$\mathbf{P}^{k+1} = \mathbf{P}^k + \mathbf{d}^k$$

We stop the algorithm when the step length is sufficiently small: $\|\mathbf{d}^k\|_2 \leq \epsilon$.

However, each iterate of the algorithm is not certified to be feasible. Indeed, for obtaining the search direction $\mathbf{d}^k$, the demand constraint is linearized at $\mathbf{P}^k$. The final solution presents only a small feasibility error because it is related to the step size $\mathbf{d}^k$ that is sufficiently small at the end of the algorithm. In order to obtain a feasible solution that is close to the optimum, the last iterate of the SQP can be projected on the feasible domain. However, in practice, it is observed that the error made on the optimal solution of the SQP is of order of 10^{-9} MW for the 100-unit problem. Even by projecting this solution on the feasible domain by using **Gurobi**, the solver provides a feasible tolerance such that the solution still has a 10^{-11} MW-order error. Therefore, taking the large run time for computing the projection into consideration, the projection step presents less benefits and is neglected.

When we analyze the performance of the SQP, it beats up every solver. Indeed, in addition to solve in 8 iterations a problem of size 100, it also takes less time than the well implemented methods of **Scipy**. The results are shown Figure 4.9.1 and it takes approximately 2 seconds for solving the iterates comparing to 6 seconds for the best method of **Scipy** that is *trust-constr* for the non-convex problem with 100 generators.

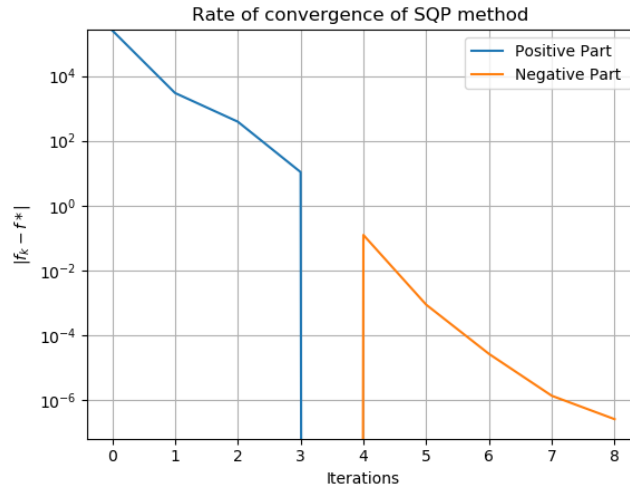


Figure 4.9.1: Convergence for the static problem with 100 generators.

From the observations made in this section, the second order solvers of **Scipy** and the SQP are very efficient for solving the static EED. Indeed, these methods are well designed for nonlinear optimization programs and are more efficient than the first-order solvers presented in the previous section. They can handle problems with more variables like the 40 and 100-unit test case and are in addition, even more efficient. The SQP algorithm is noted to be the best suited program. However, the SQP works well on this problem because it contains an unique nonlinear constraint. Indeed, additional nonlinear constraints that might be considered cannot be linearized causing the risk of being incompatible together leading to a lack of convergence in the method.

Chapter 5

Prohibited Operating Zones

The last chapter of this thesis adds another important component to the static Economic Emission Dispatch problem: the Prohibited Operating Zones (POZ) of the generators. This makes the problem very interesting since the POZ results in integer variables in the model. Mixed Integer Nonlinear Programming (MINLP) that includes integer and continuous variables, presents the algorithmic challenges of combinatorial and nonlinear programming. The input of decision variables contributes therefore to a non-convex domain leading to a non-convex POF for this multi-objective problem. This chapter presents therefore the analysis and some results on the static problem considering the POZ.

The first section describes the importance of the POZ and presents the modifications that it brings to the model. Then, Section 5.2 tackles the multi-objective problem using the scalarization method. It is interesting to perform this analysis in order to localize the non-convex segments of the front. Moreover, it analyzes also the the outputs of each generators and explains the shape of the front that has to be expected.

Finally, the last section analyzes the non-convex regions of the front with the ϵ -constraint method. Some observations and conclusions are made with the complete POF on the static EED using the POZ.

5.1 Problem formulation

It is common that power plants cannot generate at all amounts of power inside their production range. Due to physical limitations of power plant components, units cannot provide a continuous range of output power generation [31]. A discontinuous input/output power generation has therefore to be considered and can be illustrated as in Figure 5.1.1.

In order to model the problem, the static version is considered with the nonlinear objective function and using the transmission losses. The permitted operating zones for each generator can be written as:

$$\begin{aligned} P_i^{min} &\leq p_i \leq P_{i,1}^l \\ P_{i,j-1}^u &\leq p_i \leq P_{i,j}^l \\ P_{i,n_i}^u &\leq p_i \leq P_i^{max} \end{aligned} \quad j = 2, \dots, n_i \quad (5.1.1)$$

where n_i represents the total number of POZ of unit i and $P_{i,j}^l$ and $P_{i,j}^u$ are the lower and upper limit of the j^{th} POZ. These disjoint regions form therefore a non-convex feasible set.

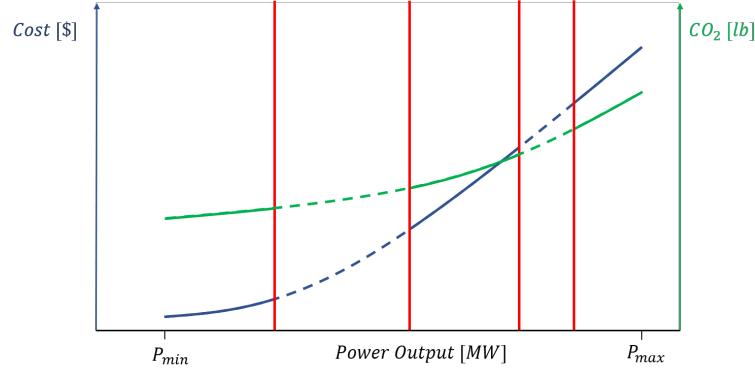


Figure 5.1.1: Unit input/output characteristic.

Since the the optimal solution has to respect one and only one constraint given in (5.1.1) for each i , binary variables are required in order to write the optimization program. The optimization problem becomes now a Mixed Integer Nonlinear Program (MINLP) and the integer constraints are:

$$\begin{aligned}
 p_i &\leq b_{i,1}P_{i,1}^l + (1 - b_{i,1})P_i^{max} & i = 1, \dots, N & \quad (5.1.2) \\
 b_{i,j}P_{i,j-1}^u &\leq p_i & j = 2, \dots, n_i \quad i = 1, \dots, N & \\
 p_i &\leq b_{i,j}P_{i,j}^l + (1 - b_{i,j})P_i^{max} & j = 2, \dots, n_i \quad i = 1, \dots, N & \\
 b_{i,n_i+1}P_{i,n_i}^u &\leq p_i & i = 1, \dots, N & \\
 \sum_{j=1}^{n_i+1} b_{i,j} &= 1 & i = 1, \dots, N & \\
 b_i &\in [0, 1]^{n_i+1} & i = 1, \dots, N &
 \end{aligned}$$

From Section 4.9, it is shown that the SQP method is the most efficient solver for this problem without the POZ. Therefore, this method is used on the new model and has to be extended by considering the constraints from (5.1.2). Recall that the SQP algorithm is described in (4.9.5). From this, the following constraints on the search direction problem (4.9.4) have to be included:

$$\begin{aligned}
 P_i^k + d_i &\leq b_{i,1}P_{i,1}^l + (1 - b_{i,1})P_i^{max} & i = 1, \dots, N & \quad (5.1.3) \\
 b_{i,j}P_{i,j-1}^u &\leq P_i^k + d_i & j = 2, \dots, n_i \quad i = 1, \dots, N & \\
 P_i^k + d_i &\leq b_{i,j}P_{i,j}^l + (1 - b_{i,j})P_i^{max} & j = 2, \dots, n_i \quad i = 1, \dots, N & \\
 b_{i,n_i+1}P_{i,n_i}^u &\leq P_i^k + d_i & i = 1, \dots, N & \\
 \sum_{j=1}^{n_i+1} b_{i,j} &= 1 & i = 1, \dots, N & \\
 b_i &\in [0, 1]^{n_i+1} & i = 1, \dots, N &
 \end{aligned}$$

5.2 Scalarization of the objective

This section presents the POF of this problem obtained by the scalarization method. From Section 1.3, it is explained that this method cannot provide all Pareto optimal values on a non-convex problem. Indeed, it reaches only partially the global front but reveals the information

that the holes hide non-convex regions of the front. These results are therefore analyzed and explained here based on the 6-unit test system.

By applying the scalarization method on the POZ model, it is expected to obtain certain regions of the POF because the feasible set is not continuous and therefore non-convex. The results for the 6-unit test case are displayed on Figure 5.2.1 and can be compared with the convex front obtained with the analog problem that does not consider the POZ (magenta on the figure). There are indeed 5 holes that cannot be attained from this method. The reason is because the gradient values at these regions correspond already at another location on the front. This means therefore that the empty regions either contain an irregular change of slope or a discontinuity in the front. Section 5.3 will indeed analyze the nature of these irregularities via the ϵ -constraint method.

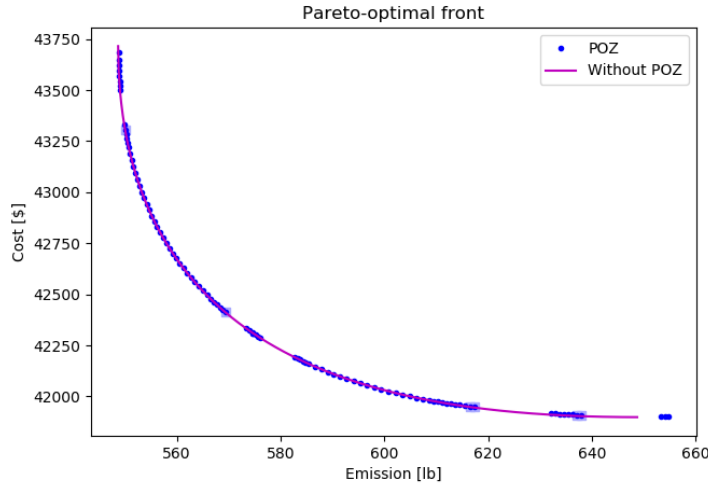


Figure 5.2.1: Attainable points on the Pareto optimal front from the scalarization technique for the 6-unit test system and $P_D = 800$ MW.

By considering these Pareto objective values, it is interesting to have a look at the Pareto solutions that are associated to these points. The outputs are displayed on Figure 5.2.2 for the different weights on the objective functions. Recall that the method for spacing equally the weights is explained in Section 1.3.2 with the average price on the emissions given from (1.3.4). Note that the grey regions on the figure correspond to the POZ of each generator. Every jump from a feasible zone to another is related to a discontinuity on the Pareto curve of the latter figure. In addition, it is also observed that some Pareto objective values correspond exactly to those for the problem with no consideration of the POZ. This happens when none of the outputs are limited by the limits on the operating zones.

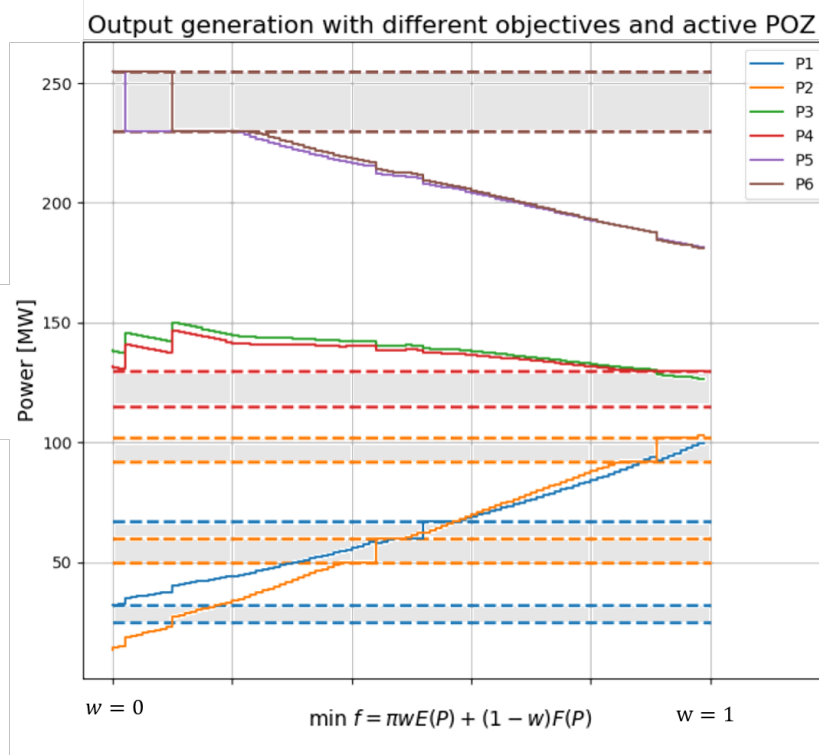


Figure 5.2.2: Output for different weights in the objective for the 6-unit system and $P_D = 800$ MW.

From these latter observations, it can be deduced that each continuous part on the front is associated to a particular set of operating zones. In order to understand the nature of the global Pareto front, an analysis of the individual sub-fronts, identified by each of these sets, can be performed. Figure 5.2.3 illustrates indeed the sub-POF's for different operating zones. The black curve, called convex front in the figure, is obtained without considering the POZ. The global POF for the problem including the POZ, can be seen as the union of these sub-fronts. It is observed that they intercept each other not smoothly, resulting to a non-smooth point on the global front. In addition, it is not always true that sub-fronts intercepts other ones as it can be noticed in the figure by the sub-front of zone 1. This results to a discontinuity in the global front between 643 and 647 lb. Indeed, on this interval, the blue curve of the first optimal front has a higher cost comparing to the lowest cost of the orange POF. On the other side, the feasible objective space is well defined in this interval because the blue sub-POF attains lower emission values than the maximum emission of the orange curve.

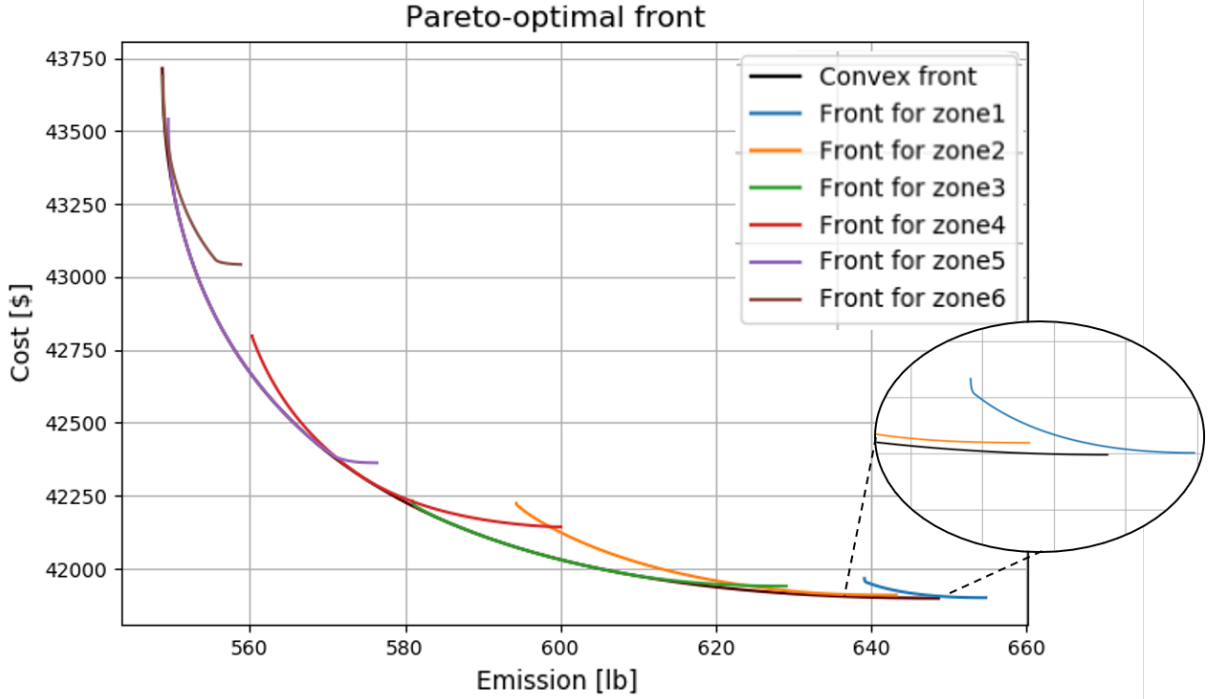


Figure 5.2.3: Pareto optimal fronts for different Operating Zones.

5.3 ϵ -constraint method

From the previous section, it is shown that the POF presents some non-convex regions that cannot be reached by the scalarization method. This section uses therefore the ϵ -constraint method for obtaining the complete front.

Since it is known from Section 4.2.2 that the ϵ -constraint method is able to reach the entire POF, it is well suited for the current model. In order to solve the considered MINLP with the ϵ -constraint, it is necessary to modify the SQP model and particularly the minimization problem associated to the search direction $\mathbf{d}$ given in (4.9.4). As observed in Section 4.9.1, additional nonlinear constraints cannot be approximated locally by a first-order approximation, otherwise there will be no more guarantee on the convergence. By considering the additional mixed integer constraints (5.1.3), the ϵ -constraint has to be added and is given by :

$$F(\mathbf{P}^k + \mathbf{d}) \leq \epsilon \quad (5.3.1)$$

$$\text{or } E(\mathbf{P}^k + \mathbf{d}) \leq \epsilon \quad (5.3.2)$$

depending on which objective we constrain.

As it is observed in the conclusion of Section 1.5, the choice of the objective function to constrain, depends on the slope of the Pareto front. Indeed, constraining the costs is more pertinent when the Pareto curve is nearly vertical and inversely, a better estimation of the front near the Economic Dispatch solution is obtained by constraining the emissions. Therefore, all Pareto optimal values on the left-side of the front are obtained by constraining the costs and the right-side is obtained by constraining the emissions. The center of the front is defined by

the cost and emission value obtained with the solution of the problem with objective:

$$f(\mathbf{p}) = \frac{\pi}{2}E(\mathbf{p}) + \frac{1}{2}F(\mathbf{p}). \quad (5.3.3)$$

On Figure 5.3.1 is displayed the complete POF. As expected from Section 5.2, the front is non-smooth and the discontinuity is noticed in the expected region. However, note that the POF is discontinuous but the feasible objective space is defined inside this interval of emissions as it was observed from Figure 5.2.3.

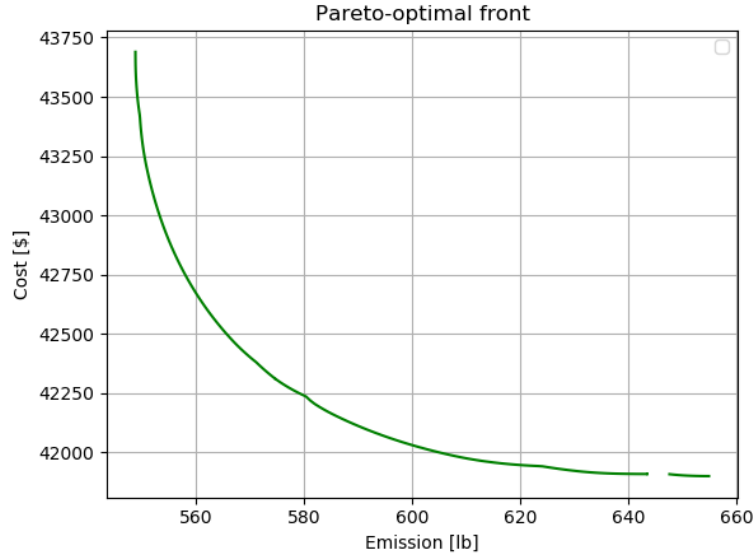


Figure 5.3.1: Pareto optimal front for the 6-unit test case with $P_D = 800$ MW.

It is now possible to measure the difference between the Pareto curve of the problem with and without the POZ. In order to compute this difference, we refer on the methods described in Section 4.7.1. The method on the w -objective is not applicable since all points on the Pareto front are not attainable via the scalarization. The first presented method, consisting in computing the error following the axes, gives a too large error and is also not very pertinent. Therefore, the second method is used and consists in comparing two Pareto optimal values at minimal distance in the Euclidean norm.

Let us consider that the reference curve is the POF without the POZ. An interesting measure is to determine the maximal difference between the reference points and the feasible objective space of the POZ problem. The feasible objective space is indeed considered instead of the POF because it is defined for all emissions and costs in the attainable range while the POF is not. As a result, the maximum difference between the two curves corresponds to a difference of 0.7 *lb* in emission rejection or 0.1% relatively and 15.55 \$ or a relative difference of 0.04% in fuel costs.

As a conclusion, the Prohibited Operating Zones of the power plants are important features for the problem since it is common that the optimal power dispatches are not physically feasible given the POZ. Therefore, it is required to modify the model in order to take this into consideration leading to a MINLP. The problem, becoming MIP, works on a discontinuous feasible set and is more difficult to solve due to this characteristic. Therefore, the POF results

in a non-smooth curve containing discontinuities. However, from the results on the 6-unit test case, the Pareto front is not strongly impacted in terms of objective values. This is of course based on a test case but gives already a valid intuition for real systems containing more units. Indeed, the more generators there are, the less the adaptations related to the POZ will be felt at the level of the Pareto frontier.

Conclusion

The Economic Emission Dispatch (EED) optimization problem is addressed and analyzed in this master's thesis. Considering that the EED problem concerns the fossil fuel fired power plant operations, it consists of allocating the power production among generators while lowering both the fuel costs and the emission level while taking the operational constraints into account. The aim of this work was (1) to develop and contribute to a deterministic algorithm for solving the problem and (2) to analyze the impact of the pollutant emissions in the problem. Since the complete model is large in dimension, containing nonlinear objective functions and a huge number of constraints, it was necessary to mathematically consider simpler formulations of the problem. By adding step by step more complex components, it was possible to improve the model.

Indeed, the first chapter tackled the simple static model with quadratic objective functions while considering generation capacity and the linear demand constraint. This Quadratic Program (QP) could be easily solved with existing commercial software tools and the notions of multi-objective optimization were introduced and illustrated on this problem.

Then, the second chapter dealt with the dynamic EED by considering the ramp rate limits and the spinning reserve constraints. It is shown that the model remains QP and that the model presents solutions that are able to handle fluctuating power demands.

Chapter 3 described a better approximation of the emission function of the generators leading to changing the complexity of the problem in a convex nonlinear program.

In addition, Chapter 4 considered the transmission losses in the system requiring to deal with a non-convex optimization problem due to the non-convex domain. Analyses were performed on the static model in order to obtain convex and linear relaxations of the constraint. It results that the convex relaxation model provides the exact solution given certain conditions of the problem's parameters. Therefore, on this relaxation, several first-order and second-order algorithms were applied in order to observe their efficiency. It turns out that the Sequential Quadratic Programming (SQP) method presented the best in-class results for the static problem using the transmission losses.

Finally, the most complex version of the model is considered in the last chapter by adding the Prohibited Operating Zones of the power plants. It resulted in a Mixed Integer Nonlinear program (MINLP) and it is showed that the SQP method can handle this problem in combination with the MIP solver of **Gurobi**.

However, the best considered model in this report, which is a MINLP, can be enhanced further more and is therefore a well-studied domain for solving it efficiently. For example, the valve-point effects on the fuel cost function or the unit commitment consideration are important features to be considered. Currently, there exist no efficient deterministic programs that can handle completely the problem, which is an opportunity for future research.

Finally, this master's thesis showed that the emissions related to fossil fueled power plants can be reduced by changing only the power distributions between the units, this being the major aim of the EED problem. This is obviously a short-term alternative to diminish the emissions by the energy sector but it has been noted in several chapters that no drastic emission decreases are to be expected in order to provide the well-desired carbon neutrality.

Appendix A

Test cases

A.1 2-unit case

Simplified version of A.3: generator n° 1 and 4 are considered.

A.2 3-unit case

Simplified version of A.3: generator n° 1, 2 and 4 are considered.

A.3 6-unit case

Mainly inspired from the IEEE 30 Bus System:

Parameters	Values
Fuel cost and emission function	From [9]
Ramp rates	From [10]
Loss matrix	From [9]
POZ	From [10]

A.4 10-unit case

All parameters are given in [11].

A.5 40-unit case

Parameters	Values
Fuel cost and emission function	From [9]
Loss matrix	Extension of A.4

A.6 100-unit case

Personal case creation:

Parameters	Values
Fuel cost and emission function	Extension of A.5
Loss matrix	Extension of A.4

Bibliography

- [1] André Jol and Hans-Martin Füssel. “Climate change, impacts and vulnerability in Europe 2016”. In: *European Environment Agency* (2017), pp. 12–30.
- [2] European Parliament. *What is carbon neutrality and how can it be achieved by 2050?* Oct 2019. URL: <https://www.europarl.europa.eu/news/en/headlines/society/20190926ST062270/what-is-carbon-neutrality-and-how-can-it-be-achieved-by-2050> (visited on 2020-06-02).
- [3] Matthieu Combe. *Les émissions mondiales de CO2 ont encore augmenté en 2019*. Dec 2019. URL: <https://www.techniques-ingenieur.fr/actualite/articles/les-emissions-mondiales-de-co2-ont-encore-augmente-en-2019-73425/> (visited on 2020-06-02).
- [4] Le Quéré et al. “Global Carbon Budget 2015”. In: *Earth Syst. Sci. Data*, (2015), pp. 12–30.
- [5] Wikipedia contributors. *Fossil fuel power station*. URL: https://en.wikipedia.org/wiki/Fossil_fuel_power_station (visited on 2020-06-02).
- [6] Liam Critchley. *How are Thermal Power Plants Polluting the Environment?* Sept 2019. URL: <https://www.azocleantech.com/article.aspx?ArticleID=861> (visited on 2020-06-07).
- [7] S. Pan, J. Jian, and L. Yang. “A hybrid MILP and IPM approach for dynamic economic dispatch with valve-point effects”. In: *International Journal of Electrical Power Energy Systems* 97 (2018), pp. 290–298.
- [8] T. Dinesh Varma and D.M. Vinod Kumar. “Multi-Objective Economic Emission Load Dispatch using Teacher-Learning- Based Optimization Technique”. In: *IFAC Proceedings Volumes* 47 (March 2014), pp. 819–826.
- [9] M. Basu. “Economic environmental dispatch using multi-objective differential evolution”. In: *Applied Soft Computing* (2011), pp. 2845–2853.
- [10] Suresh K. Damodaran and T. K. Sunil Kumar. “Combined Economic and Emission Dispatch Using a Classical Technique”. In: *IFAC Proceedings Volumes* 47 (March 2014), pp. 1049–1053.
- [11] Guesmi and Tawfik. “Extended Dynamic Economic Environmental Dispatch using Multi-Objective Particle Swarm Optimization”. In: *International Journal on Electrical Engineering and Informatics* 8 (March 2016), pp. 117–131.
- [12] Stephen Boyd and Lieven Vandenbergh. *Convex optimization*. Cambridge, UK: Cambridge University Press, 2009, pp. 174–188.
- [13] A. Oukennou, A. Sandali, and S. Elmoumen. “Coordinated Placement and Setting of FACTS in Electrical Network based on Kalai-smorodinsky Bargaining Solution and Voltage Deviation Index”. In: *International Journal of Electrical and Computer Engineering* 8.6 (Dec 2018), pp. 4079–4088.

- [14] Endre Süli and David F. Mayers. *An Introduction to Numerical Analysis*. Cambridge, UK: Cambridge University Press, 2003, pp. 300–302.
- [15] RTE. *Daily load curves*. May 2020. URL: <https://www.services-rte.com/en/view-data-published-by-rte/daily-load-curves.html> (visited on 2020-05-27).
- [16] C. Wang and S.M. Shahidehpour. “Effects of ramp-rate limits on unit commitment and economic dispatch”. In: *IEEE Transactions on Power Systems* Vol.8.3 (1993), pp. 1341–1350.
- [17] FUERGY. *The Duck Curve: Why Renewables Still Rely on Fossil Fuel Backups*. Sept. 2018. URL: <https://www.fuergy.com/cz/blog/the-duck-curve-why-renewables-still-rely-on-fossil-fuel-backups> (visited on 2020-05-18).
- [18] Hannah Ritchie and Max Roser. “Renewable Energy”. In: *Our World in Data* (2020). <https://ourworldindata.org/renewable-energy>. (Visited on 2020-05-27).
- [19] Anthony Papavasiliou. *Optimization Models in Electricity Markets*. UCLouvain, 2016.
- [20] T. Niknam et al. “A New Modified Teaching-Learning Algorithm for Reserve Constrained Dynamic Economic Dispatch”. In: *IEEE Transactions on Power Systems* Vol.28.2 (May 2013), pp. 749–763.
- [21] M.R. Gent and J.W. Lamont. “Minimum emission dispatch”. In: *IEEE Transactions on Power Apparatus and Systems* PAS-90 (1971), pp. 2650–2660.
- [22] LLC Gurobi Optimization. *Function constraints, quadratic constraints*. Feb 2020. URL: <https://www.gurobi.com/documentation/9.0/refman/constraints.html%5C#%20%5C%5Csubsection:QuadraticConstraints>.
- [23] S. Sudhoff. *Lecture 9: Multi-Objective Optimization*. URL: <https://engineering.purdue.edu/~sudhoff/ee630/Lecture09.pdf> (visited on 2020-05-27).
- [24] Yu. Nesterov. *Introductory Lectures on Convex Programming : Basic course*. 1998.
- [25] Wikipedia contributors. *Matrix representation of conic sections*. Oct 2019. URL: https://en.wikipedia.org/wiki/Matrix_representation_of_conic_sections (visited on 2020-05-22).
- [26] L. Van Hoorebeeck and P. -. A Absil and A. Papavasiliou. “Tight Bounds on Non-convex Dispatches”. preprint on webpage at perso.uclouvain.be/loic.vanhoorebeeck/abstracts/TightBounds.html. 2020.
- [27] Yiu-Wing Leung and Yuping Wang. “U-measure: a quality measure for multiobjective programming”. In: *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 3.33 (May 2003), pp. 337–343.
- [28] E.G. Birgin, J.M. Martinez, and M. Raydan. “Spectral Projected Gradient methods: Review and Perspectives”. In: *Journal of statistical software* 60.3 (November 2012), pp. 1–21.
- [29] The SciPy community. *Optimization (scipy.optimize)*. Dec 2019. URL: <https://docs.scipy.org/doc/scipy/reference/tutorial/optimize.html>.
- [30] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. New York, NY: Springer Series in Operations Research and Financial Engineering, 2006, pp. 529–562.
- [31] Fred N. Lee and Arthur M. Breipohl. “Reserve Constrained Economic Dispatch with Prohibited Operating Zones”. In: *IEEE Transactions on Power Systems* Vol.8.1 (1993), pp. 246–254.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/epl