UCLouvain

École polytechnique de Louvain

Convex hull pricing in local energy markets

Author: **Alexandre De Cannière**
Supervisor: **Anthony Papavasiliou**
Readers: **Dirk Roose, Mehdi Madani**
Academic year 2019–2020
Master [120] in Mathematical Engineering

Abstract

The increasing decentralization of power production in electricity networks (linked to the increasing share of distributed renewable resources, such as solar panels) sets up new challenges and, among them, the need for coordination between the transmission system operator and the distribution system operator. The potential of this coordination offers not only the opportunity to optimize the grid balance at a broader level, but also the opportunity to optimize the choice of the prices more globally, which is the subject of this work.

While basing our model on an second order cone (SOC) relaxation of the optimal power flow (OPF) problem, a specific dual function optimization method is studied (namely, the "level method") for the purpose of the computation of one of the pricing scheme (namely, the convex hull price) and different pricing approaches for real and reactive power are then analyzed on an actual, country-scale dataset (provided by SmartNet).

Our analysis reveals that the level method should be disregarded when the size of the dual space is large but that the convex hull price is still of interest since an approximate of it can give a lower uplift than other pricing schemes. More broadly, on the analyzed scenarii, the uplifts required are negligible compared to the profits generated.

Acknowledgements

My first thank goes to my supervisor, Professor Anthony Papavasiliou for his time, sincerity and for his very contagious passion for the electrical world that boosted me at every step of this work.

Secondly, I am very grateful for the precious hand Ilyès Mezghani provided me. His inputs were extremely valuable and allowed me to successfully conduct this analysis.

I also want to express my profound gratitude to Professor Dirk Roose and to Mehdi Madani for agreeing to be part of this thesis jury.

Last but not least, I would like to thank my family for its support throughout these five years and to thank my better half for her appreciable expertise in optimization and her love.

Contents

List of Notations

Sets

TN	Set of nodes of the transmission network.
DN	Set of nodes of the distribution networks.
N = $(TN) \cup (DN)$	Each node belongs either to the transmission network or to the distribution networks
L_T	Set of lines of the transmission network.
L_D	Set of lines of the distribution networks.
$L = (L_T) \cup (L_D)$	Each line belongs either to the transmission network or to the distribution networks, the line between the T&D belongs to the T.
$T = \{1...t_{final}\}$	Time horizon and time steps of the market clearing problem.
$SB/QB/QtB$	Respectively the set of segment bids, Q bids and Qt bids.
MDP	Set of *minimium duration pairs*. Pairs of Q bids such that the second should be activated if the first one is.
$ExQB/ExQtB$	Set of exclusive options for Q bids and Qt bids. Only one Q bid / Qt bid can be activated for the associated set.

Parameters

S_l	Power limit of line l $\in L$.
P_{sb}	Real power quantity of segment bid sb $\in SB$.
$\Delta P_n/\Delta Q_n$	The real/reactive power demand at node n $\in N$.
$\underline{x_{sb}}/\overline{x_{sb}}$	Lower/upper limit on the fraction of activation of $sb \in SB$.
B_l	Susceptance of transmission line l $\in L$.
R_l	Resistance of line l $\in L$.
X_l	Reactance of line l $\in L$.
G_l	Conductance of line l $\in L$.
$\underline{RP}/\overline{RP}$	Lower/upper ramp production bound.
$V^n_{min/max}$	Voltage bounds at a node $n \in DN$.
A_n	Unique ancestor of node n $\in DN$ (since the distribution networks are acyclic).
C_n	Children of node n $\in DN$ (since the distribution networks are acyclic).
$priceLow/priceHigh$	Price associated to the first/last MW bid (this is only of importance in curtailable bids, otherwise $pLow = pHigh$).

Variables

p_n	Real power production at node n $\in$ N.
q_n	Reactive power production at node n $\in$ N.
$fp_l/fp_{n_1n_2}$	Flow of real power of line l $\in$ L / from node 1 to node 2, $n_1, n_2 \in N$.
$fq_l/fq_{n_1n_2}$	Flow of reactive power of line l $\in$ L / from node 1 to node 2, $n_1, n_2 \in N$.
x_{sb}	Fraction of activation of segment bid sb $\in$ SB.
s_{sb}	Activation of segment bid sb $\in$ SB.
q_{qb}	Activation of Q-bid qb $\in$ QB.
qt_{qtb}	Activation of Qt-bid qtb $\in$ QtB.
θ_n	Phase angles at each node.
csq_l	Current squared over line l.
vsq_n	Voltage squared at node n.
α	binary var, 1 if the q-bid starts at time t, 0 otherwise.
ω	binary var, 1 if the q-bid finishes at time t, 0 otherwise.

List of Figures

List of Tables

Chapter 1

Literature review

1.1 Motivation: The T&D coordination problem

In a conventional centralized electric grid, the transmission network (TN) "transmits" the high-voltage power from power plants to the distribution network (DN) that "distributes" it to the consumers. Nowadays, with the increasing share of distributed renewable resources, such as solar panels, we are moving towards a decentralized network where the power is generated in the distribution network as well. However, so far, the loads of the distribution system are aggregated at a transmission node and the optimization of the grid balance takes place at the transmission level.

Using the power coming from the distribution network to the benefit of both the transmission network operator (TSO) and the distribution network operator (DSO), would help to achieve the renewable energy integration targets set worldwide by policy makers. However, in order to achieve a reliable and safe energy supply while maximizing the account on renewable energy (which is by nature uncertain), a strong TSO-DSO coordination is needed. This subject gets currently more and more attention from the academic world and practitioners [2], [7], [14].

Coordination schemes won't be investigated in this work, instead we will consider a TSO-DSO operator (that we will call the "network operator") managing the whole network. Even though this scenario is unlikely to happen, it provides a benchmark for further investigations.

In this work we will focus on the *pricing* of electricity. Although this subject has received a lot of attention, both from the practitioners and from the academic world, long before the willingness to better coordinate TSO and DSO appeared (e.g.

[3], [9], [16]), the potential of this coordination offers, not only the opportunity to optimize the grid balance at a broader level, but also the opportunity to optimize the choice of the prices more globally. This last point is the core of the present document.

This work has three goals:

- To briefly expose the mathematical problem of grid management, the state of the art pricing schemes and promising dual function optimization methods.
- To study a specific dual function optimization method (namely, the "level method") for the purpose of the computation of one of the pricing scheme (namely, the convex hull price).
- To analyze different pricing schemes for real and reactive power on an actual, country-scale dataset.

The first objective will be addressed in this first chapter. The second chapter will introduce the specific model we are working with. Finally, the third and fourth chapters will tackle respectively the second and third items of the goal list.

1.2 OPF problem and simplifications

In this section, we will start by covering the problem this work seeks to tackle -the optimal power flow problem (OPF)- and we will discuss the need for a simplification of this problem. Then we will introduce and derive the choices that have been made in this thesis, namely a linearization in the case of the transmission network and a second-order cone convexification in the case of the distribution network. This section will stay close to the excellent summary given by J. Taylor in [20].

1.2.1 The Optimal Power Flow problem

An electrical grid is basically a graph of nodes and lines, where nodes can take or inject power and lines allow power to flow between nodes. While managing an electrical grid, finding the best dispatch of power between producers and consumers while respecting the network constraints is known as the Optimal Power Flow (OPF) problem.

A first question that arises is "What does best mean?". Several answers exist. Classically the objective can be to minimize the operating costs, to minimize the losses over the network or to **maximize the social welfare** (see figure 1.1) which

will be our choice in this work.

Let us now introduce it mathematically.
If we consider a set of generators G and a set of loads L, we can characterize the generators by an integrable increasing marginal cost function:

$$MC_g : \mathbb{R} \to \mathbb{R} \quad \forall g \in G$$

and the loads by an integrable decreasing marginal benefit function:

$$MB_l : \mathbb{R} \to \mathbb{R} \quad \forall l \in L.$$

Let p_g be the production of generator g and d_l be the demand of load l, we obtain the following expression of the welfare:

$$\sum_{l \in L} \int_0^{d_l} MB_l(x)dx - \sum_{g \in G} \int_0^{p_g} MC_g(x)dx.$$

A second logical question is "How challenging is it to respect the network constraints?". The principles of the physics of electricity are accurately described by the famous Kirchhoff voltage and current laws. However, these laws create non-linearities and nonconvexity, making the OPF problem rather challenging to solve. Approximations of this problem are then needed in practice and the rest of this section will clarify that point.

Basic formulation

Complex voltage is classically used to formulate the OPF problem and, in what follows, v_i denotes the complex voltage at node i. If a pair of nodes (i,j) is linked by a line, the real and reactive power flows between this pair is p_{ij} and q_{ij}. Let p_i^p denotes the real power produced at node i and p_i^d denotes the real power demand at node i, the same notation is used for reactive power as well with respectively q_i^p and q_i^d.
A complex impedance is defined on each line $Z_{ij} = R_{ij} + iX_{ij}$ and an admittance (which is the inverse) is also defined $Y_{ij} = G_{ij} + iB_{ij}$. Every line has also a power capacity S_{ij}.

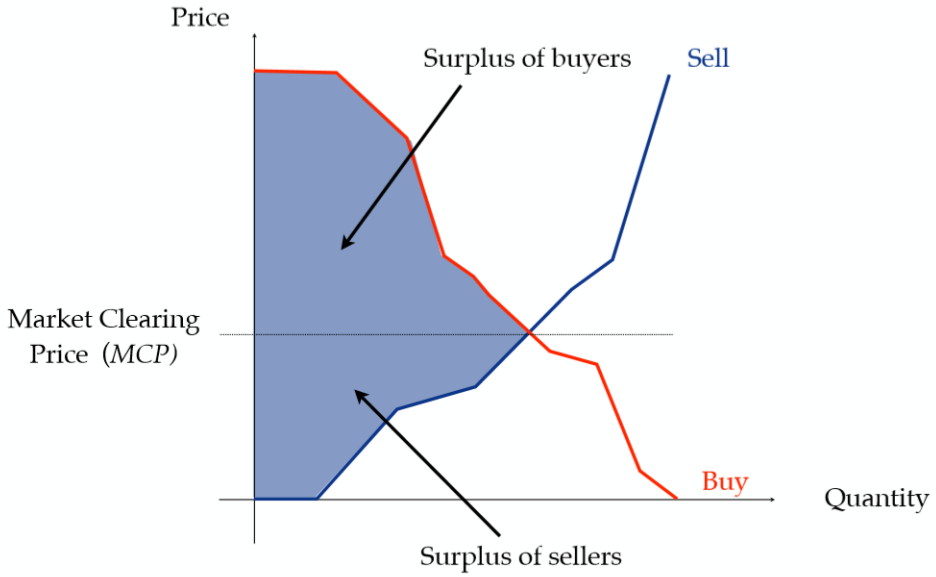

Figure 1.1: Graphical illustration of social welfare, the lines represent the aggregation of the MC/MB functions (from [5]).

The model expresses itself as

$$\max_{p,q,v} \quad \text{welfare}(p,q,v) \tag{1.1}$$

$$\text{subject to} \quad p_{ij} + iq_{ij} = v_i(v_i^* - v_j^*)Y_{ij}^* \tag{1.2}$$

$$\sum_j p_{ij} = p_i^p - p_i^d \tag{1.3}$$

$$\sum_j q_{ij} = q_i^p - q_i^d \tag{1.4}$$

$$\underline{p_i^{p/d}} \leq p_i^{p/d} \leq \overline{p_i^{p/d}} \tag{1.5}$$

$$\underline{q_i^{p/d}} \leq q_i^{p/d} \leq \overline{q_i^{p/d}} \tag{1.6}$$

$$p_{ij}^2 + q_{ij}^2 \leq S_{ij}^2 \tag{1.7}$$

$$\underline{v_i} \leq |v_i| \leq \overline{v_i} \tag{1.8}$$

(1.2) is the relationship between power flow and voltage. (1.3)-(1.4) express the net powers into or out of node i as the sum of power on the lines connected to i. (1.5)-(1.6) correspond to box constraints on the generation output/demand input range. (1.7) expresses the limit on apparent power (which is a quadratic convex constraint). (1.8) expresses the voltage magnitudes limits at each node.

Because of constraint (1.2), this model is a quadratically constrained nonconvex problem (QCP) meaning it is **NP-hard** (i.e. there exists no efficient algorithm for exactly solving it), which is why (together with their usually large size) OPF problems are particularly difficult to solve. It should also be noted that the model presented here does not provide a perfect modelization of the physical process. Dynamics, harmonics and transformer tap positions are for instance omitted. However, since we are already dealing with an NP-hard problem, going further in the modelization refining process would be unrealistic in the view of obtaining exploitable solutions.

The rest of this section is dedicated at finding a solvable simplification of this "basic formulation".

1.2.2 Simplification for the transmission network

Starting from the Kirchhoff's current and voltage laws and expressing them in terms of polar coordinates (where $v = |v|e^{i\theta}$) we obtain

$$p_{ij} = G_{ij}|v_i|^2 - |v_i||v_j|(G_{ij}\cos(\theta_i - \theta_j) - B_{ij}\sin(\theta_i - \theta_j)) \tag{1.9}$$
$$q_{ij} = B_{ij}|v_i|^2 - |v_i||v_j|(G_{ij}\sin(\theta_i - \theta_j) - B_{ij}\cos(\theta_i - \theta_j)) \tag{1.10}$$

Thanks to the four following approximations we can obtain the linearized power flow:

1. $g_{ij} = 0$: Line conductance is negligible.
2. $|v_i| = 1$: All voltage magnitudes are close to 1 per unit.
3. $\sin(\theta_i - \theta_j) = \theta_i - \theta_j$: Voltage angle differences are small.
4. Reactive power is negligible in comparison with real power. All reactive power variables and constraints can therefore be neglected.

These approximations can be justified by the *"the minor role of losses, the reduced significance of reactive power flows, the less significant role of voltage constraints, and a number of other technical factors"* [14]. More information on that matter can be found in [1].

Linearized power flow Under these assumptions, we obtain

$$\begin{aligned} & p_{ij} = B_{ij}(\theta_i - \theta_j) \\ & \sum_j p_{ij} = p_i^p - p_i^d \\ & \underline{p_i^{p/d}} \leq p_i^{p/d} \leq \overline{p_i^{p/d}} \\ & |p_{ij}| \leq S_{ij} \end{aligned} \tag{1.11}$$

which is the linearized power flow (also referred as "DC OPF" for the link that can be made with the Ohm's law).

1.2.3 Simplification for the distribution network

For a long time the kind of approximations presented above was the only way to convexify the problem. However when looking at the distribution network, neglecting the losses and the couplings between real and reactive power is not acceptable anymore. Instead of relying on approximations, we will have to use another tool called the *relaxation of constraints*, which basically "lifts" constraints to make the problem easier to solve. Of course, we do not want to create unfeasible solutions which will be a concern throughout this document.

As explained, relaxation of constraints makes possible to "jump" from nonconvex QCP to more gentle optimization classes -such as semidefinite programming (SDP) or second order cone programming (SOCP)- while taking into account the non-linearity of power flows. Thanks to the emergence of interior point methods since the mid-80's, solving theses classes of problems in a reasonable amount of time is now possible. Different relaxed models have been proposed over the years and the purpose of this work is certainly not to present them all, the interested reader can refer himself to [20] for a very complete treatment of that matter. In the following of this section, only the choice that has been made for this work will be presented. It is a SOC relaxation called the **branch flow model** which has been widely used (e.g. among others [6], [11] and [14]). Specifically, in [14] the branch flow model is tested against other relaxed models for the same first dataset that will be analyzed (the vanilla case presented in chapter 2) and other SmartNet-related datasets that will not be investigated in this document. The results were encouraging so pursuing in this direction seemed promising.

The branch flow model

The branch flow model for radial (acyclic) network was presented for the first time in 1989 by M. Baran and F. Wu [12]. Since it will be a corner stone for the rest of this work, its complete derivation is presented here.

Starting from the complex power loss from node i to node j:

$$|I_{ij}|^2 \; Z_{ij} = \frac{p_{ij} + q_{ij}}{|v_i|^2} \; Z_{ij}$$

Replacing $|I_{ij}|^2$ and $|v_i|^2$ by csq and vsq for the sake of conciseness, we obtain:

$$csq \; Z_{ij} = \frac{p_{ij} + q_{ij}}{vsq} \; Z_{ij}$$

The convention followed is that the departing power is positive and the arriving one is negative, assuming R_{ij} and $X_{ij} > 0$, the line losses are positive and equal to $p_{ij} + p_{ji}$ and $q_{ij} + q_{ji}$. Hence, we obtain:

$$p_{ij}^2 + q_{ij}^2 = csq \; vsq \tag{1.12}$$
$$p_{ij} + p_{ji} = R_{ij} \; csq \tag{1.13}$$
$$q_{ij} + q_{ji} = X_{ij} \; csq \tag{1.14}$$

The voltage at node j is given by Ohm's law:

$$v_j = v_i - I_{ij} \; Z_{ij} \tag{1.15}$$

Equations (1.12)-(1.15) together with power conservation and line limits offer an exact description of power flow, equivalent to (1.2)-(1.8).

Convexification From equation 1.15, taking the squared complex magnitude:

$$vsq_j = |v_i - I_{ij} \; Z_{ij}|^2 \tag{1.16}$$
$$= vsq_i - 2Re\{vsq_i^* \; I_{ij} \; z_{ij}\} + |I_{ij}|^2 \; |Z_{ij}|^2 \tag{1.17}$$
$$= vsq_i - 2(R_{ij} \; p_{ij} + X_{ij} \; q_{ij}) + (R_{ij}^2 + X_{ij}^2) \; csq_{ij} \tag{1.18}$$

The last source of nonconvexity arises from (1.12). Switching from an equality to the inequality "$\leq$" changes it into an hyperbolic constraint which is among the variety of the second order cone (SOC) constraints.
Obviously, in order to obtain physically meaningful solutions, **we want this constraint to be tight**. Even though it can sound optimistic, experimental and theoretical results suggest that, in the case of radial networks, this relaxation leads to high quality results [6], [14]. This element will be part of our observations when discussing the results in chapter 4.

SOCP relaxation of OPF Ultimately, the model obtained expresses itself in the following form (taking back the notations from the "basic formulation"):

$$\begin{aligned}
\max_{p,q,v} \quad & \text{welfare}(p,q,v) && (1.19)\\
\text{subject to} \quad & p_{ij}^2 + q_{ij}^2 \leq csq\ vsq && (1.20)\\
& p_{ij} + p_{ji} = R_{ij} I_{ij}^2 && (1.21)\\
& q_{ij} + q_{ji} = X_{ij} I_{ij}^2 && (1.22)\\
& v_j^2 = v_i^2 - 2(R_{ij} p_{ij} + X_{ij} q_{ij}) + (R_{ij}^2 + X_{ij}^2) I_{ij}^2 && (1.23)\\
& \sum_j p_{ij} = p_i^p - p_i^d && (1.24)\\
& \sum_j q_{ij} = q_i^p - q_i^d && (1.25)\\
& \underline{p_i^{p/d}} \leq p_i^{p/d} \leq \overline{p_i^{p/d}} && (1.26)\\
& \underline{q_i^{p/d}} \leq q_i^{p/d} \leq \overline{q_i^{p/d}} && (1.27)\\
& p_{ij}^2 + q_{ij}^2 \leq S_{ij}^2 && (1.28)\\
& \underline{v_i} \leq |v_i| \leq \overline{v_i} && (1.29)
\end{aligned}$$

1.3 Electricity pricing

The goal of this section is to address the difficulties that electricity pricing raises and to present the methods already proposed to solve these issues.

Electrical energy is traded in an electricity market which is a *"centralized platform where participants can exchange electricity transparently according to the price they are willing to pay or receive, and according to the capacity of the electrical network"* [5]. The main advantage of this market is its capability to break monopolies and open the producers and retailers of electricity to competition but, to do so, it requires electricity prices to balance transactions.
In fact, two different electricity markets live side by side: the day-ahead market (where participants can buy or sell electricity for the next 24 hours in a closed auction) and the intra-day market (where participants can continuously bid and where, when a deal is feasible, it is executed). In the present work, only the **day-ahead market** will be considered.

Before jumping into the details of the computation of the price, the question of "where" should be asked. Indeed, as already mentioned, the electrical network is composed of nodes and lines and the way to model the power over the lines has been extensively described in the previous sections. One way of considering the pricing states that each node can be associated to a different price (this is called "**nodal pricing**" or "locational marginal pricing") and this approach will be followed in the present work.

However, it should be noted that, in Europe, each country wishes to have a uniform price over its territory; even though their territories contain more than one node. Nodes are then aggregated at a national level, the prices are computed and a re-dispatch of the electricity is performed to ensure that the lines' constraints are respected.

1.3.1 Pricing in the convex case

In order to simplify the discussion here, let us introduce a simpler model than the OPF model (1.2)-(1.8) called the **economic dispatch model** (1.30). Indeed, the network's constraints are not needed to highlight the concepts presented here. This section is inspired by [17].

In this model, we neglect the reactive power, and we keep the notations introduced before, namely: G and L sets of generators and loads, $MC_g : \mathbb{R} \to \mathbb{R}, \ \forall g \in G$ and $MB_l : \mathbb{R} \to \mathbb{R}, \ \forall l \in L$ integrable increasing/decreasing marginal cost/benefit function and p_g/d_l the production/consumption of an agent. The under-brackets variables represents the dual variables of the constraints.

$$\begin{aligned}
&\max_{p_g d_l} \sum_{l \in L} \int_0^{d_l} MB_l(x)dx - \sum_{g \in G} \int_0^{p_g} MC_g(x)dx \\
\text{subject to} \quad &\sum_{g \in G} p_g - \sum_{l \in L} d_l \geq 0 \quad : (\lambda) \\
&p_g \leq \overline{p_g} \quad : (\mu_g) && \forall g \\
&d_l \leq \overline{d_l} \quad : (\mu_l) && \forall l \\
&p_g \geq 0 && \forall g \\
&d_l \geq 0 && \forall l
\end{aligned} \tag{1.30}$$

Since all the constraints are linear and the objective is convex, **the problem is convex**. The constraint associated to the dual variable λ is called the **power balance constraint** or the market clearing constraint since it makes the link

between production and demand and ensure the equality. Hence, using duality theory, the dual variable λ represents a variation of the total cost of the system when the demand is perturbed. The Karush Kuhn Tucker conditions (KKT conditions) are the classical mathematical tool used for analyzing dual variables. Applied to (1.30) these conditions become:

$$0 \leq \mu_g \perp \overline{p_g} - p_g \geq 0 \tag{1.31}$$
$$0 \leq \mu_l \perp \overline{d_l} - d_l \geq 0 \tag{1.32}$$
$$0 \leq p_g \perp \mu_g - \lambda + MC_g(p_g) \geq 0 \tag{1.33}$$
$$0 \leq d_l \perp \mu_l + \lambda - MB_l(d_l) \geq 0 \tag{1.34}$$
$$0 \leq \lambda \perp \sum_{g \in G} p_g - \sum_{l \in L} d_l \geq 0 \tag{1.35}$$

Proposition 1.3.1. *Given an optimal solution of the economic dispatch problem, there exists a threshold λ such that:*

1. *If a generator is operating strictly within its dispatch interval $(0 < p_g < \overline{p_g})$, then $MC_g(p_g) = \lambda$. If a load is consuming strictly within its dispatch interval $(0 < d_l < \overline{d_l})$, then $MB_l(d_l) = \lambda$.*

2. *If a generator is producing zero, then $MC_g(p_g) \geq \lambda$. If a load is consuming zero, then $MB_l(d_l) \leq \lambda$.*

3. *If a generator is producing at peak capacity, then $MC_g(p_g) \leq \lambda$. If a load is consuming at peak capacity, then $MB_l(d_l) \geq \lambda$.*

Proof. The results comes from the KKT conditions of the problem. For instance, suppose that a generator g produces between 0 and $\overline{p_g}$. From (1.31), we have that $\mu_g = 0$ and from (1.33) we have that $\lambda = MC_g(p_g) + \mu_g = MC_g(p_g)$. The other statements follow the same logic. □

This is economically called marginal pricing which means that the production cost of the plant which is at the margin of production is chosen as the price.

1.3.2 Pricing with nonconvexities

As chapter 2 will show for our specific case, the economic dispatch model is in general nonconvex. The nonconvexities can arise from different sources but a classical source used in *example 1* is the addition of fixed costs, in fact, adding fixed costs as well as other nonconvexity sources will often require the use of **discrete**

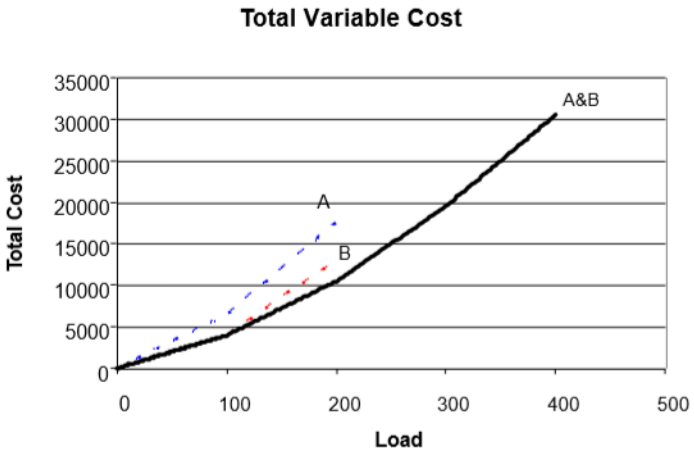

Figure 1.2: Aggregate cost illustration - no fixed costs (from [16]).

variables.

The nonconvexity of the model is a bad news since it means that strong duality does not hold any longer. Hence, the challenge to define a totally new pricing rule is faced. Before presenting the potential new pricing rules, an example is introduced to give the reader a bit of intuition on the challenges faced. This example was firstly presented by P. Gribik in [16].

Example 1: A two plants example.
Consider two generation units with the parameters presented in the table below.

	q(MW)	Plants A	B
Fixed Cost (€)		0	6000
Var Cost 1 (€/MWh)	100	65	40
Var Cost 2 (€/MWh)	100	110	90

At first, we do not consider the fixed costs. On figure 1.3 one can see the evolution of the marginal price (i.e. the variable cost) of the MWh as a function of the load. For the first 100MW, the cheapest plant is B at 40€/MWh, then A is the cheapest for the 100 following MW and so on. Following the discussion in section 1.3.1, for any level of load, one can easily find on this figure the marginal prices that will become the market clearing prices (i.e. a price such that the producers have the incentives to commit and to produce such that the market clears). One

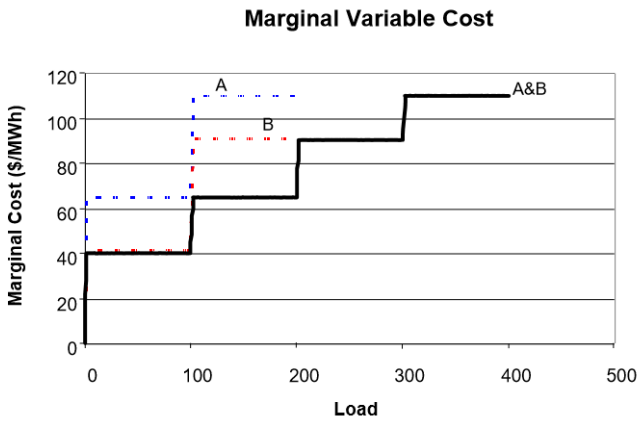

Figure 1.3: Marginal cost illustration - no fixed costs (from [16]).

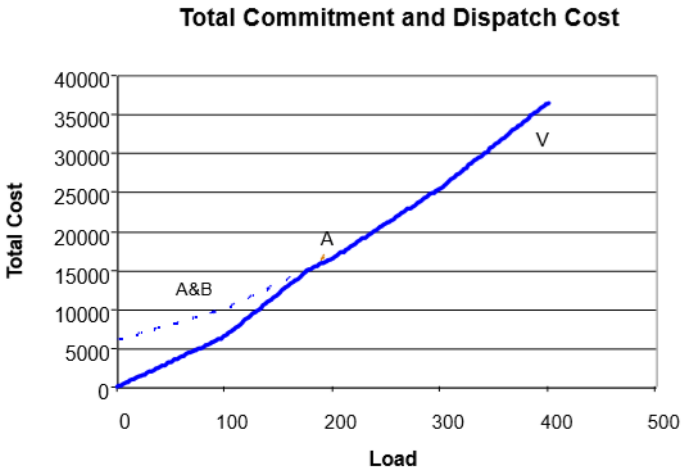

Figure 1.4: Aggregate cost illustration - with fixed costs (from [16]).

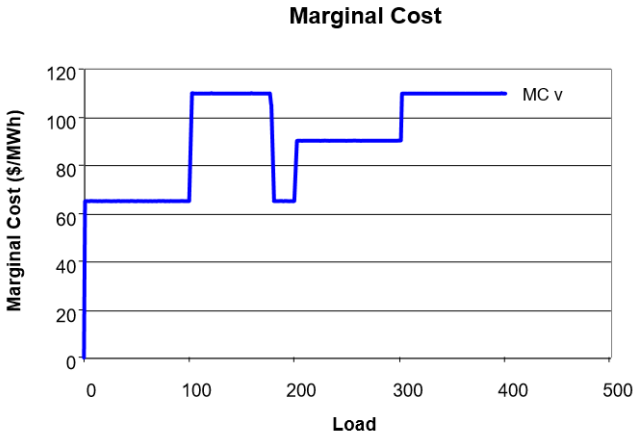

Figure 1.5: Marginal cost illustration - with fixed costs (from [16]).

should note that the prices are increasing with the load, which seems natural.

Now let us add the fixed costs. Figure 1.5 looks now quite different from figure 1.3, the marginal cost does not constantly increase with the load anymore but increases then decreases then increases again. The existence of a market clearing price is not guaranteed anymore when the convexity is lost and the question of **"how to create an efficient price nevertheless?"** rises.

Different approaches have been developed to answer this question and the contribution of this work is to study their behavior on a real, country-scale, T&D network example (see chapters 3 and 4). For now on, they will be presented from a theoretical point of view[1].

Problem (1.36) formulates the economic dispatch problem in a very schematic way while separating the continuous variables (x) from the discrete ones (u). Note that u^* denotes the optimal commitment.

Schematic problem

$$\begin{aligned} \max_{x,u} \quad & w(x,u) \\ s.t. \quad & g(x) = y \\ & u \in \{0,1\} \end{aligned} \tag{1.36}$$

[1]This summary of the existing approaches is inspired by a workshop given by W. Hogan at UCLA in January 2016 [9].

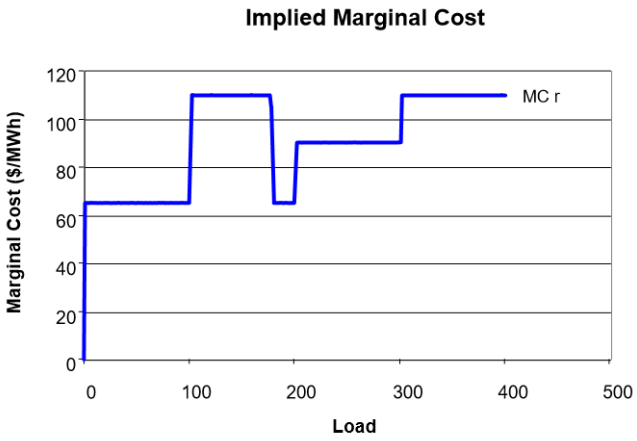

Figure 1.6: Marginal cost illustration - Restricted approach (from [16]).

Restricted Model

Proposed in 2005 by O'Neil [3], the idea is here to fix the unit commitment variables (at an optimal solution). With those variables fixed, the model is now convex and it is possible to determine the marginal prices. Problem 1.37 sketches it out and figure 1.6 illustrates this method.

$$\begin{aligned} \max_{x,u} \quad & w(x,u) \\ s.t. \quad & g(x) = y \\ & u = u^* \end{aligned} \tag{1.37}$$

Dispatchable Model

In this scheme, even if we know that the problem is not convex, we make it convex by relaxing the discrete constraints. The marginal prices associated to this relaxed, continuous and convex problem are then taken. Figures 1.7 and 1.8 and problem 1.38 illustrate this method.

It should be noted that this method is actually already in use in some parts of the USA by `PJM Interconnection LCC`, which is an American TSO that operates over multiple states [10].

$$\begin{aligned} \max_{x,u} \quad & w(x,u) \\ s.t. \quad & g(x) = y \\ & 0 \leq u \leq 1 \end{aligned} \tag{1.38}$$

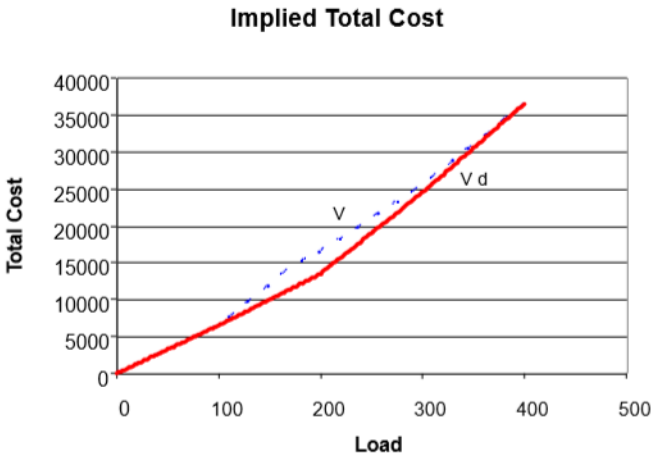

Figure 1.7: Aggregate cost illustration - dispatchable approach (from [16]).

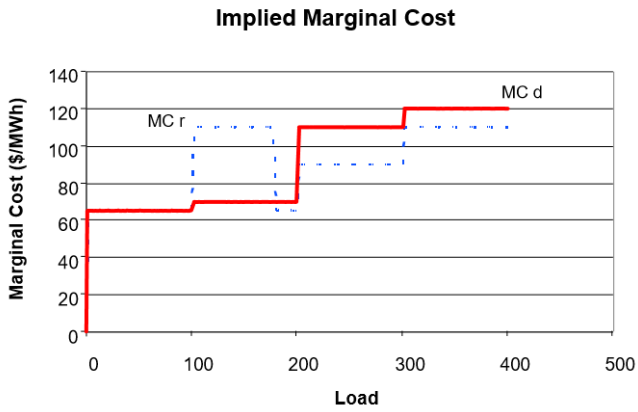

Figure 1.8: Marginal cost illustration - dispatchable approach (from [16]).

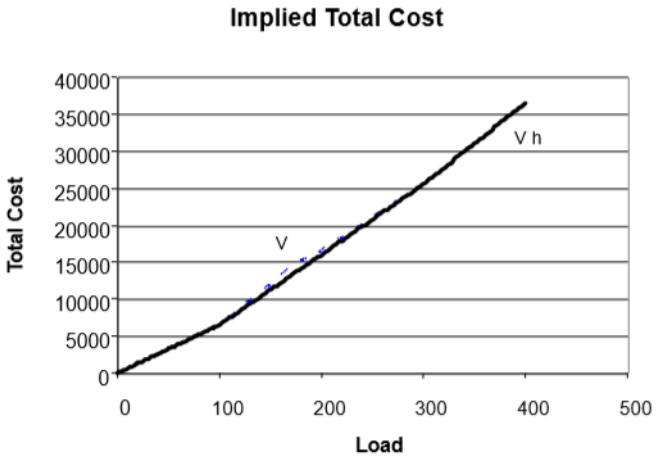

Figure 1.9: Aggregate cost illustration - CHP approach (from [16]).

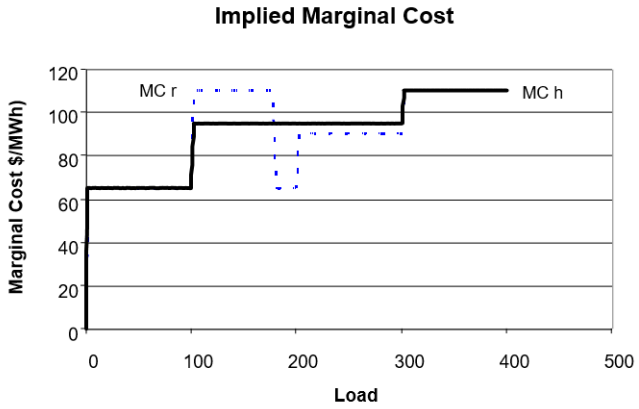

Figure 1.10: Marginal cost illustration - CHP approach (from [16]).

Convex hull pricing

Proposed by Hogan, and described in [16], the idea of convex hull pricing (CHP) is to determine the prices from the convex hull of the objective function (see figure 1.9). Or, equivalently, to determine them from the Lagrangian relaxation of the market clearing constraints (the dual variables associated are then the prices we are looking for).

As for the dispatchable approach, the CHP is also already used in the USA. The `Midcontinent Independent System Operator` already uses the CHP and it

is under consideration by `PJM Interconnection LCC`[2].

Mathematically, from 1.30, while relaxing the market clearing constraint, we find the Lagrangian function:

$$L(p_g, d_l, \lambda) = \sum_{l \in L} \int_0^{d_l} MB_l(x)dx - \sum_{g \in G} \int_0^{p_g} MC_g(x)dx + \lambda(\sum_{g \in G} p_g - \sum_{l \in L} d_l) \quad (1.39)$$

While fixing the Lagrangian multiplier λ to a certain value, we obtain the Lagrangian dual function:

$$g(\lambda) = \max_{p_g, d_l} L(p_g, d_l, \lambda) \quad (1.40)$$

Proposition 1.3.2. *$g(\lambda)$ is convex following λ.*

Proof. Let us fix p_g^* and d_l^* as the optimal reaction to $\alpha\lambda_1 + (1-\alpha)\lambda_2$, with $\alpha \in [0,1]$ and $\forall \lambda_{1,2} \in \mathbb{R}$. We obtain:

$$\begin{aligned}
&L(\alpha\lambda_1 + (1-\alpha)\lambda_2, p_g^*, d_l^*) \\
&= \sum_{l \in L} \int_0^{d_l^*} MB_l(x)dx - \sum_{g \in G} \int_0^{p_g^*} MC_g(x)dx + (\alpha\lambda_1 + (1-\alpha)\lambda_2)(\sum_{g \in G} p_g^* - \sum_{l \in L} d_l^*) \\
&= \alpha \left(\sum_{l \in L} \int_0^{d_l^*} MB_l(x)dx - \sum_{g \in G} \int_0^{p_g^*} MC_g(x)dx + \lambda_1(\sum_{g \in G} p_g^* - \sum_{l \in L} d_l^*) \right) \\
&\qquad + (1-\alpha) \left(\sum_{l \in L} \int_0^{d_l^*} MB_l(x)dx - \sum_{g \in G} \int_0^{p_g^*} MC_g(x)dx + \lambda_2(\sum_{g \in G} p_g^* - \sum_{l \in L} d_l^*) \right) \\
&\leq \alpha \left(\max_{p_g, d_l} \sum_{l \in L} \int_0^{d_l} MB_l(x)dx - \sum_{g \in G} \int_0^{p_g} MC_g(x)dx + \lambda_1(\sum_{g \in G} p_g - \sum_{l \in L} d_l) \right) \\
&\qquad (1-\alpha) \left(\max_{p_g, d_l} \sum_{l \in L} \int_0^{d_l} MB_l(x)dx - \sum_{g \in G} \int_0^{p_g} MC_g(x)dx + \lambda_2(\sum_{g \in G} p_g - \sum_{l \in L} d_l) \right) \\
&= \alpha L(\lambda_1, p_g^*, d_l^*) + (1-\alpha) L(\lambda_2, p_g^*, d_l^*)
\end{aligned}$$

□

By convex theory, we know that any local minimum of a convex function is a global minimum of this function. We have that $g(\lambda)$ is defined $\forall \lambda \in \mathbb{R}$ since λ is

[2]More information can be found at [18]

not present in the constraint of the relaxed economic dispatch model.
Hence, there exists at least one **global minimum**, but the convex hull price **might not be unique**.

The following results are central and justify all the efforts made to compute the CHP. A formal proof of these statements can be found in [16].

Proposition 1.3.3. *Convex hull price minimizes the uplift payments.*

Proposition 1.3.4. *For convex problems, convex hull pricing is equivalent to marginal pricing.*

Proposition 1.3.3 introduces the concept of *uplift payment*. Since this concept is a cornerstone of our analysis, the next paragraph will describe it in more in details.

Uplift/side payments

Starting from a proposed market clearing price p, the revenues generated from fulfilling a demand l will be pl and the costs of meeting this demand will be $C(l)$. The profits will then have the following form:

$$\pi(p, l) = pl - C(l).$$

From a generator point of view, given a price p in a competitive market, their own interest would be to maximize their profits by solving:

$$\pi^* = \max_k \{pk - C(k)\}.$$

When the solution is $k^* = l$, then p is a profit-maximization solution and no uplift payment is needed.

However, if it is not the case then *"in order to support the solution, the uplift payment would have to make the market participants indifferent between the proposed solution and the unconstrained profit"* [16]. This leads us to the definition of uplift:

$$\text{Uplift}(p, l) = \pi^*(p) - \pi(p, l),$$

which is the difference between the optimal profit given a proposed price and the constrained profit at this price.

The computation of the uplift for any pricing scheme is summarized in figure 1.11 and figure 1.12 offers a comparison of the three presented approaches in terms of uplift payment.

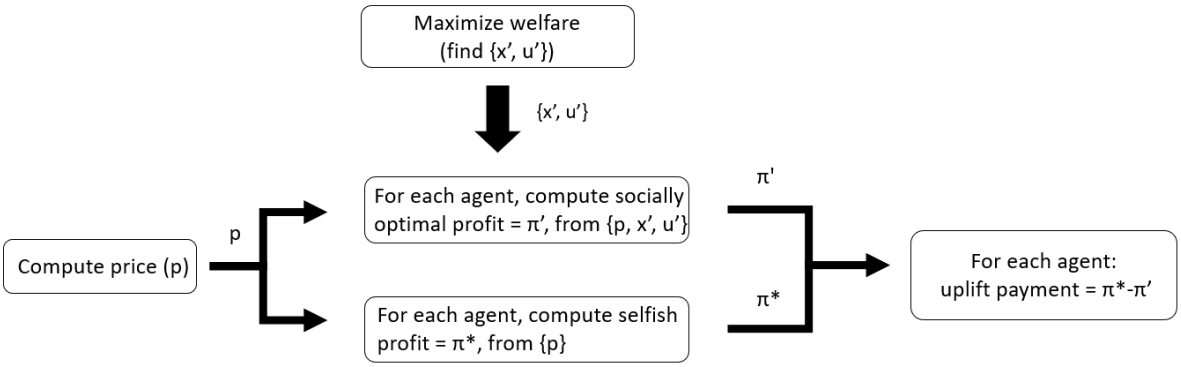

Figure 1.11: Uplift payment mechanism for a given pricing scheme. x represents the continuous variables and u the binary variables.

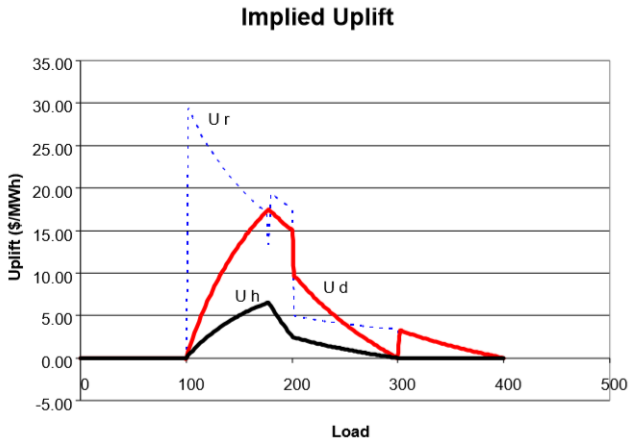

Figure 1.12: Uplift comparison between approaches (from [16]).

Uplift payments might seem not fully transparent which is why, in Europe for instance, most of the countries are reluctant to use the approaches presented above. How electricity is actually priced in Europe is briefly explained in the following section.

Since the uplift payments are a big political concern, assessing how much uplift each of these methods creates is essential. This will be one of our major goal in chapter 4.

The European case

Since 2014, for most of European countries[3] the pricing method used is an algorithm called `EUPHEMIA`. This algorithm offers the advantage to avoid uplift payment and to provide a uniform price for every country (which is required as highlighted in the introduction of this section), but those properties come at a cost.
Without entering into the details of this algorithm, to achieve these properties, the maximization of the welfare is subject to extra constraints, leading to potentially less welfare. Hence, the interest in new pricing schemes and the need of assessing the impact that the uplift would have on the extra welfare (the last point is beyond the scope of this work).

1.4 Dual function optimization

As explained in the latter section, the convex hull pricing (CHP) can be computed in two ways: a primal approach and a dual approach. The primal approach requires the explicit construction of the convex hull of the aggregate cost function, which has proven to be difficult in practice. Hence, in this work, we will focus on the computation of the dual approach, which is simpler to formulate, but harder to solve. In that view, the aim of this section is to give a theoretical summary of the first order oracle methods used here to solve dual function optimization. In chapter 3, we will show how they can be applied to our specific problem.

We will present two methods:

- The **subgradient method**: a classical and very popular easy to implement method.
- The **level method**: a probably less famous method but which has given encouraging results for computing the CHP [19].

Figure 1.13 sketches the global idea behind these methods. Note that applying the level method to the Italian scenario to observe its properties on a real, country-scale dataset is one of the goals of this thesis, as presented in section 1.1.

Let us consider the following, very general, minimization problem.

[3] 25 European countries including: Austria, Belgium, Czech Republic, Croatia, Denmark, Estonia, Finland, France, Germany, Hungary, Italy, Ireland, Latvia, Lithuania, Luxembourg, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and UK.

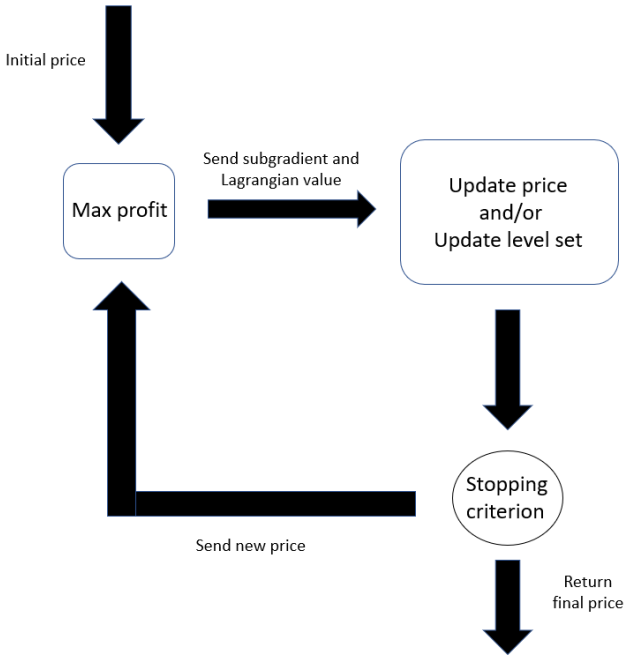

Figure 1.13: General scheme used to solve the dual optimization problem.

$$\begin{aligned} \max_{x} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \qquad i = 1..m \\ & x \in Q \subseteq \mathbb{R}^n \end{aligned} \tag{1.41}$$

Where Q is a closed, non-empty, convex set and $g_i(x)$, i = 1,..,m are convex functions, potentially nonsmooth (i.e. nondifferentiable). We can express the dual function $h(\nu)$ as

$$h(\nu) = \max_{x \in Q} \left\{ f(x) - \nu^T g(x) \right\} \tag{1.42}$$

where ν are the Lagrangian multipliers. "Relaxing" the constraints makes the problem easier to evaluate.

1.4.1 Subgradient method

Let us start with the definition of the subgradient.

Definition 1.4.1. *A vector π is called the subgradient of h at $\nu \in dom(h)$ (i.e. $\pi \in \partial h(\nu)$), if for any $\nu_0 \in dom(h)$, we have*

$$h(\nu) \geq h(\nu_0) + \langle \pi, \nu - \nu_0 \rangle$$

The subgradient method is the classical and simplest method used to solve *nonsmooth optimization* problems. It is a recursive algorithm where each new guess is computed with

$$u^{(k+1)} = u^{(k)} + \alpha_k \pi^{(k)},$$

with

- $u^{(k)}$ the k^{th} iterate
- $\pi^{(k)}$ any subgradient of h at $u^{(k)}$
- α_k the k^{th} step size

It should be stressed that the subgradient method is **not a descent method**. So, as we have no guarantee that the next iterate will be better than the previous one and the method can oscillate, we will keep track of the best iterate so far. The theoretical convergence rate of this method until the optimum is **sub-linear** but the method is not affected by the variable space size.

Step size

We will mainly use the **Polyak step**. The Polyak step is optimal when the solution is known in advance. It expresses itself as

$$\alpha_k = \frac{h(\nu^{(k)}) - h^*}{||\pi^{(k)}||_2^2},$$

with $h^* = h(\nu^*)$ the optimal solution. A proof of this result can be found in [4].

In practice, the optimal solution will of course not be known in advance, but an estimate of the solution can be used.

1.4.2 Level method

As proved in [15], the subgradient method is the theoretical optimal method for solving (1.41) when the size of the variable space is large. In practice, however,this result does not necessarily mean much since we can classically hope that the real performance of a method will be much better than the theoretical one coming from a worst case scenario.

Specifically, this is not true for the subgradient method which will **not** converge faster than in theory because of its very strict scheme. The idea of developing more flexible schemes (that could then be more efficient in practice) has lead to various new algorithms, among which the *level method* which has, as said above, already

shown to be suited for this kind of optimization.

The notion of *model* is central is these new algorithms (note that the index k denote the k^{th} iterate).

Definition 1.4.2. *Let Q be the initial domain of the problem, let $P = \{\nu_k\}_0^\infty$ be a sequence in Q and let π_k be a subgradient of h at ν_k. Then*

$$\hat{h}(\nu, k) = \min_{0 \leq i \leq k} \ [h(\nu_i) + \langle \pi_i, \nu - \nu_i \rangle]$$

denotes a model of $h(\nu)$.

Theorem 1.4.1. *Let $\hat{h}(\nu, k)$ be the model function such as defined at 1.4.2. Then $\forall k$, we have*

$$h(\nu) \geq \hat{h}(\nu, k).$$

Thus, we have that each iterate of the model function $\hat{h}(\nu, k)$ gives a **lower approximate** of our Lagrangian function $h(\nu)$.

From the model function we can define the following problem:

Model minimization

$$\begin{aligned} &\min_{\nu \in Q, \theta} \theta \\ \text{such that} \quad &h(\nu_i) + \langle \pi_i, \nu - \nu_i \rangle \leq \theta \qquad i = 1...k \end{aligned}$$

The improvement of the level method (in comparison with other *model* based methods) is its ability to update the prices more carefully.
We let $LB = \theta$ and $UB = h(\nu_k^*)$ (which denotes the smallest value of h already found) and we define $\hat{h}(\nu, k) \leq \alpha UB + (1-\alpha)LB$ to be the *level set*, which contains "better prices" than the current one with respect to the model function. The new price ν_{k+1} is then determined by the projection of ν_k onto this *level set*. The projection is expressed as a quadratic programming problem in the following way:

Projection problem

$$\begin{aligned} &\min_{\nu \in Q} ||\nu - \nu_k||_2^2 \\ \text{such that,} \quad &l_k(\alpha) \geq \hat{h}(\nu_i) + \langle \pi_i, \pi - \nu_i \rangle \quad \forall i = 1...k \\ &\text{where } l_k(\alpha) = \alpha UB_k + (1-\alpha)LB_k, \ \forall k \text{ and } \alpha \in [0, 1]. \end{aligned}$$

It should be stressed that both the projection problem and the model minimization can be solved efficiently by a simplex-type or by an interior point method. Y. Nesterov also proves in [15] that a level set of a convex function is convex and a that a projection on a convex set exists and is unique.

It should be stressed that a big advantage of the level method over the subgradient method is its ability to assess the quality of the current solution thanks to the computation of a lower and upper bound at each iteration.

Convergence properties

As a last comment on the level method, the theoretical estimated efficiency of the method (i.e. the smallest possible number of iteration before that $UB_i - LB_i \leq \epsilon$ in *worst case* scenario) is proportional to the squared diameter[4] of the variable space. This could of course lead to trouble when applying this method onto a real large examples (and as showed in section 3, it will); but once more, a theoretical upper limit could be deeply pessimistic.

There are also other reasons to be optimistic about this method. First of all, we can never know the true convergence rate of a specific problem before having tested it thoroughly. Secondly, it is sometimes possible to speed-up a method when insights about the function are known, which will be the case here since different approaches will be tested. Last but not least, in the problem we are concerned about here, we do not need a full convergence of the algorithm; we only need to obtain a "sufficiently good" solution.

The level method can be summarized as follow:

[4]The diameter of a set is defined as the upper limit of the distance between two arbitrary points of the set.

Algorithm 1: Level method

Initialization: Choose a point $\nu_0 \in Q$, accuracy $\epsilon > 0$, a level coefficient $\alpha \in [0, 1]$ and an iteration limit $iterLimit$.
Set k := 0
while $k < iterLimit$ **do**
  solve (1.42) and return π_k.
  if $(h(\nu_k) < UB)$
    then $UB := h(\nu_k)$
  solve "Model minimization" and return θ_k.
  $LB := \theta_k$
  if $(UB - LB < \epsilon)$
    then BREAK
  $l := \alpha UB + (1 - \alpha)LB$
  solve "Projection" and return ν_{k+1}
  k := k+1
end

Chapter 2

Model

This chapter presents the complete model we implemented. We start by considering the general topology of the network and the structure of the bids, and we finish with the model itself.

2.1 The SmartNet problem

The experiments exposed in chapters 3 and 4 have been performed on the Italian network thanks to data provided by SmartNet. SmartNet is a project funded by the EU that aims at *"providing architectures for optimized interaction between TSOs and DSOs in managing the exchange of information for monitoring and for the acquisition of ancillary services (reserve and balancing, voltage regulation, congestion management) both at national level and in a cross-border context"* [21].

2.2 Case studies

The datasets we will deal with contain a classical national transmission and distribution network with a transmission network (TN) connecting several radial (acyclic) distribution networks (DNs) (figure 2.2 gives schematic view of it). The topology of the network, the physical parameters of the line and the bids at each node are known.
There are two case studies we will exploit. The first one will be called the *vanilla case* and is in fact a toy example created from the second one: the *real case*. Their

Figure 2.1: The SmartNet logo

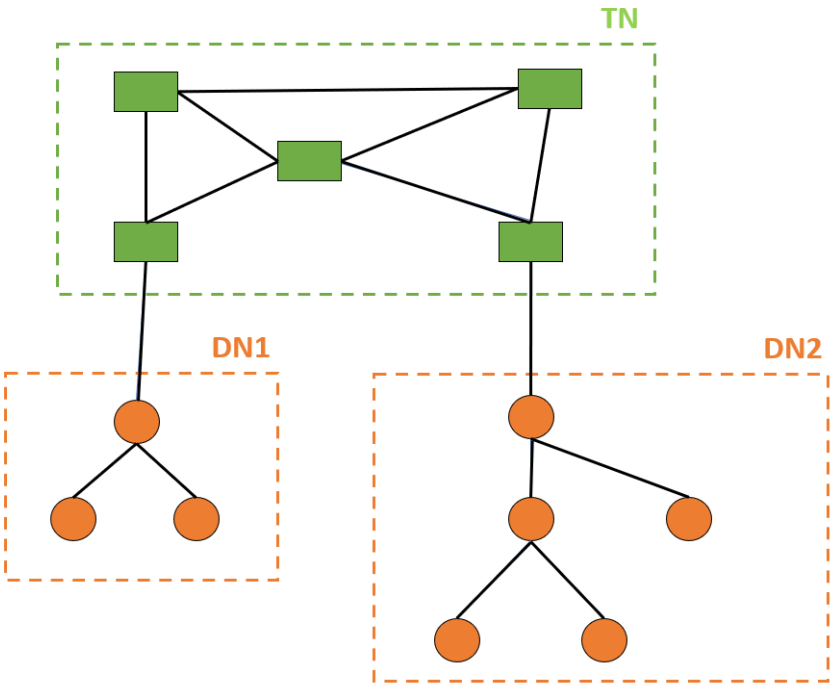

Figure 2.2: Schematic transmission & distribution network.

main characteristics are given in table 2.1. A formal definition of the bids is given in section 2.2.1. The time step is the time interval between two clearings of the market, it will have an impact on the inter-temporal constraints described below. In our case, this time step is 15 minutes.

Case study	Vanilla case	Real case
# T. nodes	27	3,648
# D. nodes	175	2,410
# DNs	4	638
# Time steps	4	3
# Unit bids	1,667	12,318

Table 2.1: Case studies overview.

2.2.1 Bids description

The products used in this work and presented here are inspired from the Central Western European (CWE) day-ahead energy market. It should be stressed that, even though the **reactive power will be priced**, there exists no bid for reactive

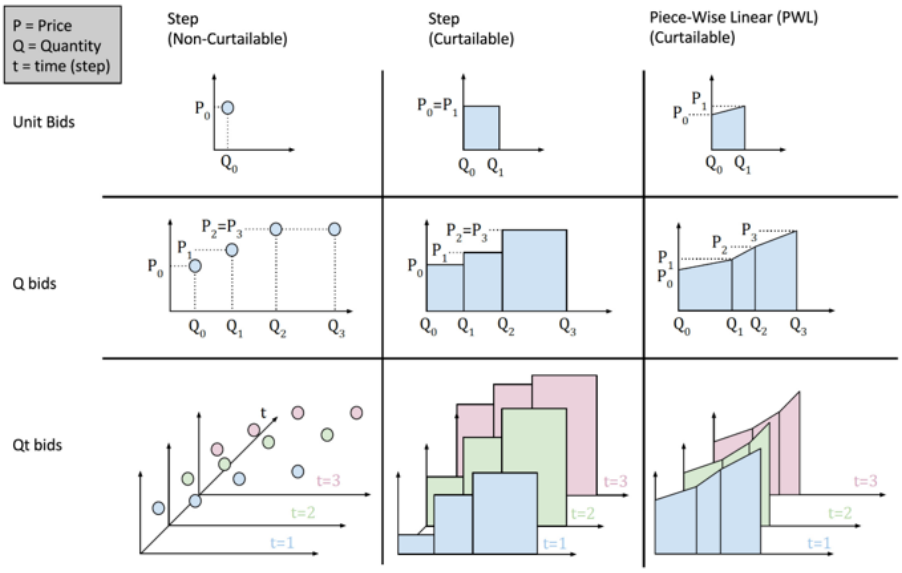

Figure 2.3: Different types of bids considered (shown only for $p, q > 0$) (from [1])

power. All the following is then only applicable for real power.
There are 3 types of bids:

- Non-curtailable unit bid : This is the simplest bid one can imagine. A price and a quantity and the operator can either accept or refuse the bid completely.
- STEP curtailable unit bid : This kind of bid is described by a single price and two quantities q_0 and q_1. Every quantity choice between q_0 and q_1 can be chosen by the operator.
- Piecewise linear curtailable unit bid : This is an extension of the step curtailable bid where the price can vary linearly with the quantity.

Unit bids (that will be referred as *segment bids* for the rest of this document) can be linked in a Q bid and Q bids can be linked over time in a Qt bid. How the bids are linked to each other is made specific in the constraints below. Figure 2.3 summarizes and graphically illustrates the whole set of possible bids.

Convention It can be noted that no assumption about the signs of the bids is done here. Hence, the terms *buyers* and *sellers* can be confusing since the bids can be negative in terms of price and in terms of quantity as well.
In a classical supply-demand market, a generator would bid with $price, quantity > 0$ and the opposite for a consumer. Nevertheless, in this less classical electricity

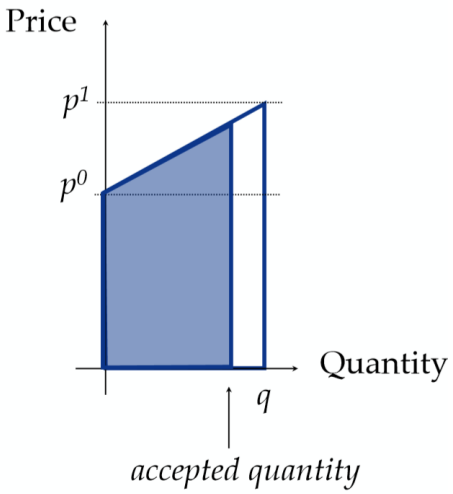

Figure 2.4: Accepted quantity illustration (from [5])

market, all different combinations are possible. For instance, a sunny and windy day may cause the network to contain more energy than predicted. In such situation, generators can be in a position where they bid a positive quantity at a negative price (they want to pay to inject energy).
The convention followed by SmartNet brings clarity about this signs issue: a bidder with $price > 0$ wants to receive money, a bidder with $price < 0$ want to pay money, a $quantity > 0$ is offered and a $quantity < 0$ is asked. The terms *buyer* and *seller* are sometimes aggregated in the generic term of *prosumer*.

Accepted quantity Every bid can be either accepted completely, partially or rejected, the variable expressing these choices is x. $x_{sb} = 0$ means that the segment bid "sb" has been rejected and all the other value up to one mean that "sb" has been accepted in a certain proportion or totally. Figure 2.4 illustrates this quantity.

2.3 Detailed model

2.3.1 Constraints

Direct current power flow for transmission

As explained in section 1.2.2, the TN is modeled using a linearized power flow often referred as "DC OPF". Using the notations presented in the list of notation at the beginning of this document, we obtain the following constraints.

$$p_n + \sum_{l=(m,n)} fp_l - \sum_{l=(n,m)} fp_l = \Delta P_n \quad ,\forall n \in TN \tag{2.1}$$

$$fp_l = B_l(\theta_n - \theta_m) \quad ,\forall l = (n,m) \in L_T \tag{2.2}$$

$$\theta_{hub} = 0 \tag{2.3}$$

$$-S_l \leq fp_l \leq S_l \quad ,\forall l \in L_T \tag{2.4}$$

$$fq_l = 0 \quad ,\forall l \in L_T \tag{2.5}$$

(2.1) represents the power balance in TN, (2.2) is the susceptance-based representation of flows, (2.3) makes the previous constraint consistent by stating that the phase angle is zero at a hub node and (2.4)-(2.5) are the line limits and the definition of reactive flow in TN (i.e. reactive flow are neglected in TN).

SOCP power flow for distribution

As explained in section 1.2.3, the DN is modeled using a second order cone relaxation of the branch flow model.

$$p_n + \sum_{j \in C_n} (fp_{jn} - csq_{jn} * R_{jn}) - fp_{nA_n} = \Delta P_n \quad ,\forall n \in DN \tag{2.6}$$

$$q_n + \sum_{j \in C_n} (fq_{jn} - csq_{jn} * X_{jn}) - fq_{nA_n} = \Delta Q_n \quad ,\forall n \in DN \tag{2.7}$$

$$(fp_l)^2 + (fq_l)^2 \leq S_l^2 \quad ,\forall l \in L_D \tag{2.8}$$

$$(fp_l)^2 + (fq_l)^2 \leq csq_l * vsq_n \quad ,\forall l = (n,m) \in L_D \tag{2.9}$$

$$vsq_n - vsq_{A_n} = 2(R_l fp_l + X_l fq_l) - csq_l(R_l^2 + X_l^2) \quad ,\forall l = (A_n, n) \in L_D \tag{2.10}$$

$$0 \leq V_{min}^n \leq vsq_n \leq V_{max}^n \quad ,\forall n \in DN \tag{2.11}$$

$$csq_i \geq 0 \quad ,\forall i \in L_D \tag{2.12}$$

(2.6)-(2.7) represents the power balance in DN (it is a simplified version of the power balance since the shunt conductance is negligible in the SmartNet dataset), (2.8) limits the apparent power, (2.9) is the SOC relaxation, (2.10) is a constraint on voltage differences and (2.11)-(2.12) are definitions of voltage and current squared.

Bid constraints

Each constraint will be shortly described just after presenting them. A more precise description can be found in the SmartNet documentation [1].

$$p_{sb}^2 + q_{sb}^2 \leq (\overline{x}_{sb} * P_{sb})^2, \quad \forall sb \in SB \tag{2.13}$$

$$p_i = \sum_{sb=(i,q,s)} P_{sb} x_{sb}, \quad \forall i \in N \tag{2.14}$$

$$\underline{x}_{sb} s_{sb} \leq x_{sb} \leq s_{sb} \overline{x}_{sb}, \quad \forall sb \in SB \tag{2.15}$$

$$s_{(i,q,s)} \leq q_{(i,q)}, \quad \forall (i,q,s) \in SB \tag{2.16}$$

$$q_{(i,qt,q)} \leq qt_{(i,qt)}, \quad \forall (i,qt,q) \in QB, (i,qt) \in QtB \tag{2.17}$$

$$qt_{(i,qt)} \leq \sum_{qb=(t,qt,q)} q_{qb}, \quad \forall (qt,i) \in QtB \tag{2.18}$$

$$\sum_{qb \in exqb} q_{qb} \leq 1, \quad \forall exqb \in ExQB \tag{2.19}$$

$$\sum_{qtb \in exqtb} qt_{qtb} \leq 1, \quad \forall exqtb \in ExQtB \tag{2.20}$$

$$q_{i,qt,q,t} - q_{i,qt,q,t-1} + \alpha_{i,qt,q,t} - \omega_{i,qt,q,t} = 0, \tag{2.21}$$

$$\alpha_{i,qt,q,t} + \omega_{i,qt,q,t} \leq 1, \tag{2.22}$$

$$q_{i,t,qt,q} \geq \alpha_{i,\tau,qt,q}, \quad \forall ((i,t,qt,q),(i,\tau,qt,q)) \in MDP \tag{2.23}$$

$$\underline{RP}_{i,t} \leq p_{i,t+1} - p_{i,t} \leq \overline{RP}_{i,t}, \quad \forall i \in N, t \in T \backslash \{t_{final}\} \tag{2.24}$$

$$0 \leq x_{sb} \leq 1, \quad sb \in SB \tag{2.25}$$

$$\alpha \in \{0,1\}^{|QB|}$$
$$\omega \in \{0,1\}^{|QB|}$$
$$s \in \{0,1\}^{|SB|}$$
$$q \in \{0,1\}^{|QB|}$$
$$qt \in \{0,1\}^{|QtB|} \tag{2.26}$$

(2.13) represents the complex injection limit, (2.14) describes the link between bids acceptance and real power injection, (2.15) defines how segment bids can be activated, (2.16) makes sure than a segment bid is activated only if the linked Q bid is also activated, (2.17) makes sure than a Q bid is activated only if the linked Qt bid is activated,(2.18 the Qt bid is activated only if at least one the associated Q bids is activated, (2.19)-(2.20) indicate that certain Q/Qt bids should be activated only if others are not, (2.21) expresses how Q bids are linked, (2.22) denotes the fact that a Q bid cannot end and start at the same moment, (2.23) imposes that a Q bid remains activated for a certain time if activated, (2.24) is a ramp constraint

on the real power output, (2.25) imposes x to be fractional and (2.26) α, ω, s, q and qt are binary variables.

We are now able to clarify the **sources of nonconvexity** assumed until now, they illustrate the lack of flexibility of thermal power plants.

- Minimum stable level: To start a power plant, a minimum range of operation is needed. The first part of equation 2.15 models this phenomenon.
- Minimum duration time: When a power plant is switched on, it should remain on for a certain amount of time (equation 2.23).
- Ramp constraint: The increase (decrease) of power output is bounded to a limited amount (equation 2.24).
- Exclusive choice: Certain bids can only be accepted if others are not (equations 2.19-2.20).

2.3.2 Objective function

The **welfare** function was already presented in chapter 1 and is the objective function of both the problems of *welfare maximization* and of *pricing* (see figure 1.11).

$$\sum_{sb \in SB} (x_{sb} P_{sb}\ priceLow + \frac{x_{sb}^2 P_{sb}(priceHigh - priceLow)}{2}) \tag{2.27}$$

The quadratic term comes from the "piecewise linear curtailable bids" presented previously (this term disappear otherwise since $priceLow = priceHigh$).

The second objective function is the **profit** function and appears in the *Profit maximization* problem of figure 1.11. Note that this function is in fact the Lagrangian relaxation of the welfare with respect to the power balance constraints (2.1)-(2.6)-(2.7). Note that (2.1)-(2.6) have been mixed to be suited for both the transmission and the distribution networks.
Since the binding constraints between the agents have been removed and each one of these agents is now **free to maximize his own profit** given the prices $\lambda^{P/Q}$.

$$\sum_{sb \in SB} \left(x_{sb} P_{sb}\ priceLow + \frac{x_{sb}^2 P_{sb}(priceHigh - priceLow)}{2} \right)$$
$$- \sum_{n \in N} \lambda_n^P \left(-p_n - \sum_{l=(m,n)} (fp_{mn} - csq_{mn}\ R_{mn}) + \sum_{l=(n,m)} fp_l + \Delta P_n \right)$$
$$- \sum_{n \in DN} \lambda_n^Q \left(-q_n - \sum_{j \in C_n} (fq_{jn} - csq_{jn}\ X_{jn}) + fq_{nA_n} + \Delta Q_n \right) \quad (2.28)$$

Chapter 3

Convex hull pricing

This section aims at clarifying how the CHP can be computed using the methods developed in section 1.4 and at testing these techniques on the case studies described in section 2.

As a reminder, the dual problem we want to solve is the following:

$$\min_{\lambda} g(\lambda) = \left\{ \begin{array}{lll} \max\limits_{p_g, d_l} \sum\limits_{l \in L} \int_0^{d_l} MB_l(x)dx - \sum\limits_{g \in G} \int_0^{p_g} MC_g(x)dx + \lambda(\sum\limits_{g \in G} p_g - \sum\limits_{l \in L} d_l) & & \\ \text{subject to: } p_g \leq \overline{p_g} & & \forall g \\ d_l \leq \overline{d_l} & & \forall l \\ p_g \geq 0 & & \forall g \\ d_l \geq 0 & & \forall l \end{array} \right\} \tag{3.1}$$

where the inner part is the selfish profit maximization.

3.1 Computation in practice

The presence of binary variables induces abrupt "switches" in the optimal dispatch decisions, hence, 3.1 will be a nonsmooth (i.e. nondifferentiable) function to optimize. Section 1.4 summarized 2 methods used to solve this difficult class of problems. These methods rely on the call to a first order oracle providing the value of the function and the value of the subgradients for a certain set of prices.
It should be noted that this call will be expensive since it requires to solve a mixed integer program.

3.1.1 Subgradient method

One of the advantages already highlighted of the subgradient method is its simplicity of implementation. The computation of the subgradients will be done according to the following proposition (a proof can be found in [8]).

Proposition 3.1.1. *Let* (p_g^*,d_l^*) *be the optimal reactions to a price* λ*. Then* $(\sum_{g\in G} p_g^* - \sum_{l\in L} d_l^*)$ *is a subgradient of* g *in* λ*.*

The choice of the step size will be made clear for each case study but the choice of the starting point will always be the dispatchable prices. The real advantage of this method for us, is its property of **not being influenced by the size of the variable space**.

3.1.2 Level method

As explained in section 1.4, the level method was developed with the idea of relying on a more flexible scheme than the strict one of the subgradient method.

Different "tricks" can be used to accelerate the convergence of the level method:

- Price box: instead of letting $\lambda \in \mathbb{R}$, we can restrict the prices in a certain range. We choose the following price range: $[-500; 3000]$, it should be noted that the same box has been taken for the real and the reactive prices.
- Starting price: Rather than starting with arbitrary prices, we can help the algorithm by giving it "good" first prices (computed with the dispatchable method for instance).
- Good level: Since this method works by projecting the prices on a level set generated by a level, helping the algorithm by giving him a "good" level from the start seems promising. If the final level is not known, we can at least approximate it.

3.2 Experimental results

3.2.1 Vanilla case study

The results of the CHP computation performed on the vanilla case study presented in chapter 2 reveal different things. Before summarizing them, let us make clear what the goal here is.

We want to compute (an approximate of) the convex hull price in the most efficient way. In this view, we will compare the relative dual gap ($\frac{\mathrm{P-W}}{|\mathrm{P}|}$) made by

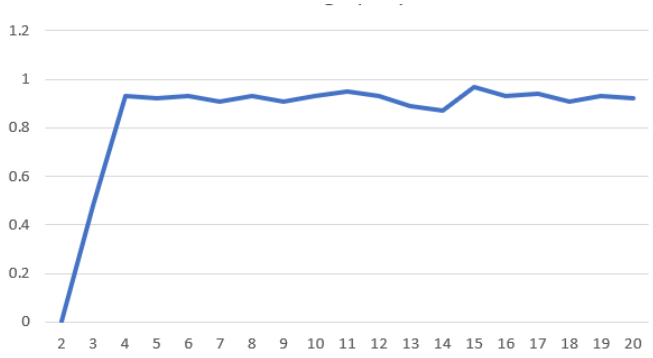

Figure 3.1: Duality gap $\frac{P-W}{|P|}$ when computing the CHP with the **"classical"** level method on the vanilla case.

each method, starting from a good starting point (e.g. the dispatchable prices). P is the selfish profit maximization and W is the maximum welfare (from weak duality, we will always have P≥W).

Since the duality gap and the uplift payments are linked (P. Gribik offers a comprehensive treatment of that matter in [16]) and since we know, from proposition 1.3.3, that the CHP minimizes these payments, the CHP must also minimizes the duality gap. Hence, in the following figures, a **smaller duality gap** that the one we started from **means less uplift payments** and prices closer to the CHP.

Figures 3.1-3.3 allow us to come to the following conclusions (the term "beat" is used as "deliver a smaller uplift"):

1. The "classical" level method (i.e. the unoptimized level method, without the tricks presented above) fails to beat the dispatchable prices after 20 iterations. In fact, the steps are far "too big" and the algorithm looses all the benefits from starting at a good point (see figure 3.1).

2. The "ameliorated" level method (i.e. optimized one, with the tricks presented above), also fails to beat the dispatchable prices after 20 iterations but remains around the initial solution (see figure 3.2).

3. The subgradient method beats the dispatchable prices. The Polyak step size or a classical constant small step size achieve globally the same (with an advantage for Polyak), see figure 3.3.

4. The duality gap is near to zero, meaning that the problem is near-convex.

The promising results we talk about when introducing the level method do not seem to be confirmed by the vanilla case study (in fact, the vanilla case study is

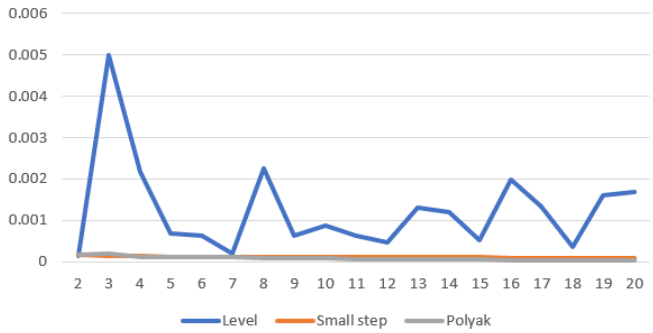

Figure 3.2: Duality gap $\frac{\text{P}-\text{W}}{|\text{P}|}$ when computing the CHP with the **"ameliorated"** level and the subgradient methods on the vanilla case.

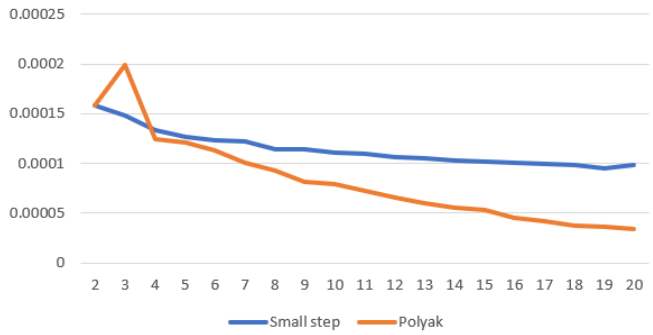

Figure 3.3: Duality gap $\frac{\text{P}-\text{W}}{|\text{P}|}$ when computing the CHP with the subgradient method on the vanilla case.

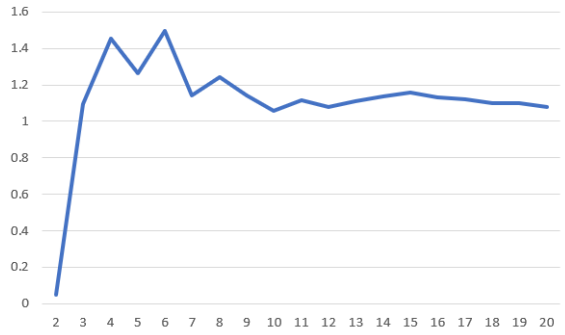

Figure 3.4: Duality gap $\frac{P-W}{|P|}$ for 20 iterations of the **"ameliorated"** level method on the real case.

already a significantly larger set than the one used in [19]). To assess whether hoping that the level method could be useful in the larger test case, trying to root source the cause of the problem would be useful.

It appears that the level method seems to suffer from size of the dual space. Nesterov in [15] proves that the theoretical estimated efficiency of the method (i.e. the smallest possible number of iteration before that $UB_i - LB_i \leq \epsilon$ in *worst case* scenario) is proportional to the squared diameter of the dual variable space (see 1.4.2).
In this example, with the prices in $[-500; 3000]^{1616}$,[1] we have $diam([-500; 3000]^{1616}) = 3500 * \sqrt{1616} \approx 140,000$.

Nesterov notes that in *real life* scenarios the convergence is usually much faster and that $\epsilon = 10^{-4} - 10^{-5}$ is obtained after $3n - 4n$ iterations (with n the size of the dual space). This result implies that, even being optimistic, for the vanilla case study, meeting a 10^{-4} precision would require roughly 5000 iterations. On a classical computer (Intel Core i5-7300HQ CPU), each iteration takes around 20 seconds, meaning that **with the level method, more than one day** of computation could be needed to meet the criterion.

3.2.2 Real case study

Our first observation will be to confirm the fact that level method does not seem suited for this kind of optimization. Here our dual space is even larger than in the vanilla case and converging in a reasonable amount of time does not seem realistic (see figure 3.4).

[1] We have 2 prices to compute at each node (real and reactive), 4 time steps and 202 nodes.

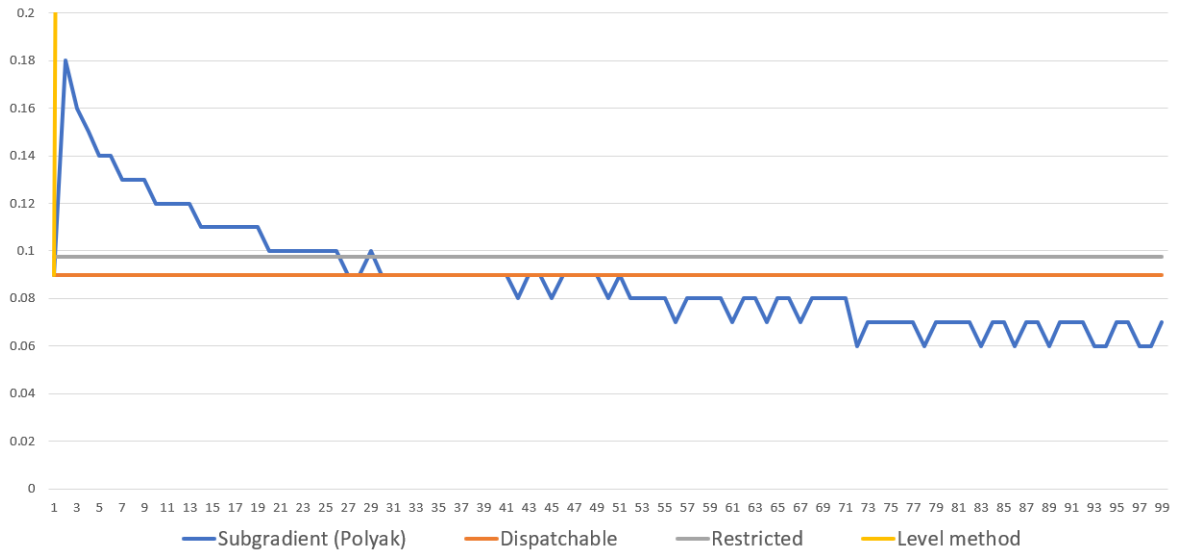

Figure 3.5: Comparison of the duality gap $\frac{P-W}{|P|}$ for the real case study.

However, the subgradient method continues to offer satisfaction. Figure 3.5 illustrates this by showing that after roughly 50 iterations the subgradient method retrieve the gap of the dispatchable solution and after 100 iterations offers around 30% of duality gap less.
It should be noted that the constant step size does not provide satisfaction anymore, its behavior is now close to the level method's behavior, even though the step size has been taken very small, the rationale behind this phenomenon could be the steepness of the function. Going a bit "too far" having disastrous consequences.

Since the prices must be computed rapidly, the time needed to compute the 100 iterations is of importance. On the same computer, the computation of a first "good" price (with dispatchable or restricted approach) takes a few minutes and every iteration of the subgradient takes around 15 seconds. 100 iterations take then around 25 minutes (without the starting price computation). For practical applications, this can then be an issue to further investigate.

To conclude, these experiments highlight the fact that **the level method should not be used** when the size of the dual space increases excessively and the best way to compute the CHP so far is to use the **subgradient with Polyak step together with a good starting point**.

Chapter 4

Economical results

This chapter aims at summarizing the main economical results obtained in this analysis. The solver used in this thesis is `Gurobi 9.0`. `Gurobi` is a commercial solver created in 2008 and is now available in various programming languages and widely used in the industry.

In the vanilla case study, our goal is to confirm the quality of the results to make sure that our analysis of the real Italian case study will be consistent.
As explained throughout this document, the amount of uplift payment is a critical quantity that we want to minimize both for its impact on the total welfare and for its political inconvenience.

However, we will also look at the slack variables of the SOC constraint (constraint 2.9). Indeed, as highlighted in section 1.2.3, for the solution to be feasible, this constraint must be tight (i.e. slacks = 0).
Finally, we will present the profits resulting of each method.

4.1 Vanilla test case

As already observed in chapter 3, the vanilla test case is near-convex and, as a consequence, the uplift should be almost zero (cf. section 1.3.2). Table 4.1 summarizes these payments and they are indeed **close to zero** over the four time steps. The slacks of the SOC constraint are equal to zero so **the solution is feasible**.

Obtaining a near-zero uplift for the vanilla case study **was not expected**. In fact, given the number of potential sources of nonconvexity quite the reverse would have been expected. In order to gain insights about this case and to better understand how "far" from a nonconvex case this scenario is, a sensitivity analysis

Vanilla case study	
Approach	Uplift payment
Restricted	0.01€
Dispatchable	0.02€
CHP	0.01€

Table 4.1: Overview of the uplift payment in the vanilla case study.

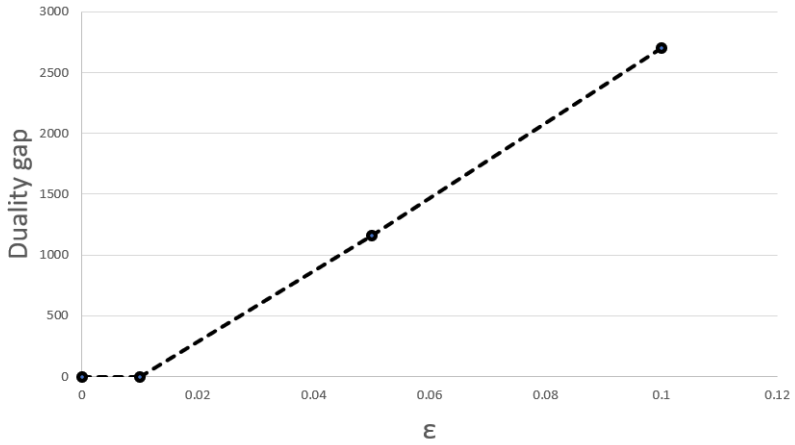

Figure 4.1: Evolution of the duality gap for $\underline{x_{sb}} > \epsilon$ on the vanilla case (dispatchable approach).

is performed (see figure 4.1). The idea is to increase the *minimum acceptable ratio* ($\underline{x_{sb}}$) which correspond to the *minimum stable level*, the first source of nonconvexity introduced in section 2.3.1.

We observe that increasing $\underline{x_{sb}}$ indeed offers a way to "deconvexify" the problem and it validates the consistency of the computation of our results. Another interesting observation is the fact that the slacks are not all equal to zero anymore; as soon as the duality gap starts to increase we observe some slacks greater than 0.

Figure 4.2 sketches the money flows between the agents, it highlights the fact that the producers earn money by selling power to the network operator (NO) -their profit is then obtained by retrieving the costs of production from theses revenues- while the NO earns money by trading power between producers and consumers. The market operator (MO) pays the uplifts to the market participants and, in the US-design, these uplifts fall back on the consumers as network fees.

In table 4.2, the profits of the producers and the NO are presented. From

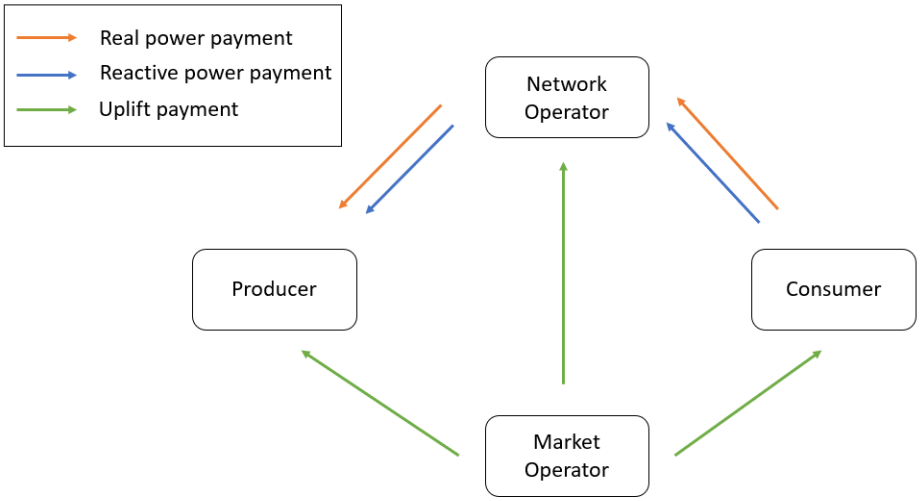

Figure 4.2: Payments flow diagram.

Profits vanilla case study	
Agent	Profit
Producers (real power)	340,147.97€
Producers (reactive power)	0€
NO (real power)	4,837.92€
NO (reactive power)	0.20€

Table 4.2: Overview of the profits made in the vanilla case study.

this table we observe that pricing the reactive power has a negligible influence on the agents' profits. Figure 4.3 illustrates that phenomenon too by showing that the reactive prices are mostly equal to zero (since the problem is near-convex, all approaches give similar prices).

4.2 Real case study

From figure 3.5, we know that, in the real case, hoping for a zero duality gap would be too optimistic, we are here in a true nonconvex case.
The uplifts per approach over the three time steps are given in table 4.3 (a visual view of it is in figure 4.4) and, interestingly, the slacks associated to the SOC constraint are all equal to zero so **the solution is physically meaningful**.

The CHP approach (approximated with 100 iterations of the subgradient

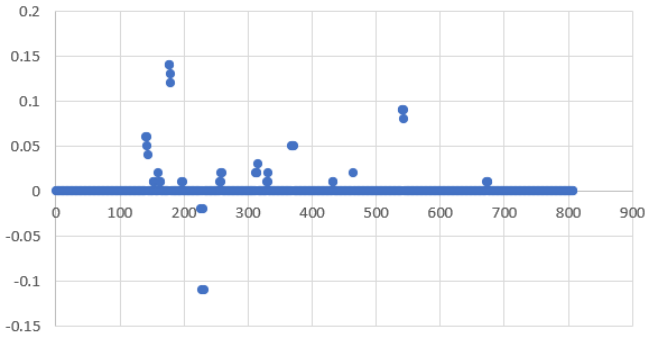

Figure 4.3: Reactive prices for every node and every time steps in vanilla case.

Real case study	
Approach	Uplift payment
Restricted	164.26€
Dispatchable	143.93€
CHP	91.44€

Table 4.3: Overview of the uplift payment in the real case study.

method) gives the lowest uplift. However, even the "worst" uplift remains very small in comparison with the total profits generated (see table 4.4). The profits are also significantly higher than in the vanilla case, simply because we have significantly more nodes and bids. We also observe that the choice of the approach has a negligible influence on the profits.

As for the vanilla case, we observe that the pricing of reactive power has a negligible influence and that the reactive prices are mostly zero (figure 4.5 illustrates this with a "box" plot where the box is a line since a vast majority of the nodes have a zero reactive price).

Profits real case study			
Agent	Dispatchable	Restricted	CHP
Producers (real power)	2,121,478.66€	2,121,490.00€	2,121,478.20€
Producers (reactive power)	0€	0€	0€
NO (real power)	60,690.86€	60,693.50€	60,690.74€
NO (reactive power)	0€	0€	0€

Table 4.4: Overview of the profits made in the real case study.

Another noticeable element comes from figure 4.6. Indeed, whatever the ap-

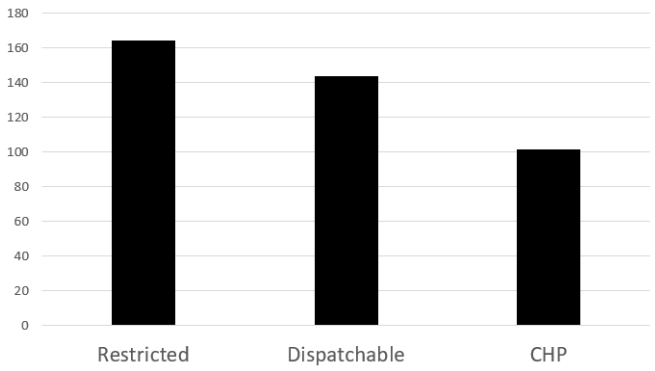

Figure 4.4: Uplift payment per method in the real case study.

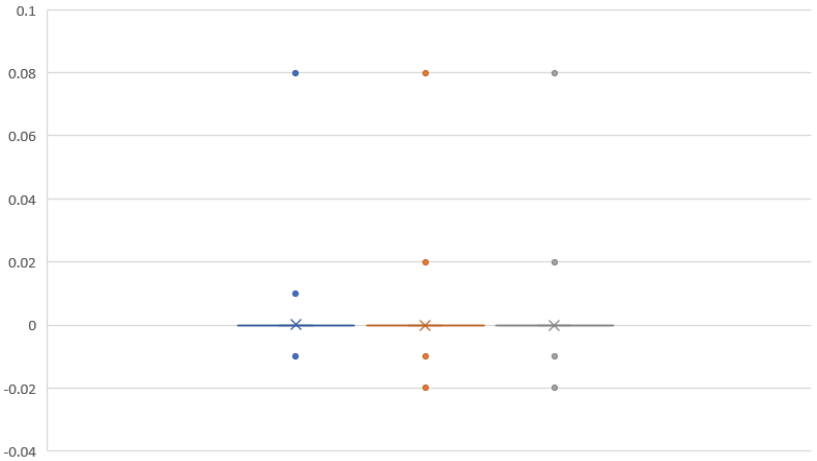

Figure 4.5: Box plot of reactive prices in real case (restricted in blue, dispatchable in orange and CHP in grey).

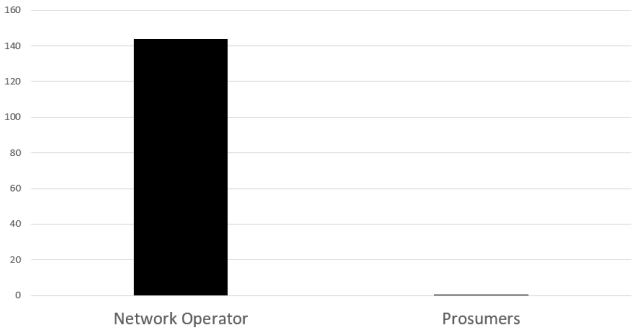

Figure 4.6: Received uplift per agent in the real case study (dispatchable approach).

proach, $> 99\%$ of the uplift is due to the NO and a the rest is due to the prosumers (where prosumers are the aggregation of the consumers and the producers). In practice, if the market operator and the NO are merged the leftover uplift would be negligible.

Chapter 5

Conclusion

The potentiality of a broader coordination between TSO and DSO offers the opportunity to optimize the choice of the nodal prices more globally.

In this work, different pricing schemes were applied to a real dataset (provided by SmartNet) containing the Italian network and a significant number of bids. The most promising pricing scheme, for its intrinsic properties, is the convex hull pricing (CHP) but unfortunately, up to now, no efficient way to compute it on large complicated dataset is known.

Recently, a new hope has arisen from the level method, which has proven to retrieve efficiently the CHP with the dual approach on a small problem. Our analysis shows however that this path seems to be a dead end while attempting to solve larger problems.

5.1 Main results

- Considering optimizing dual functions to retrieve the CHP, the level method should be disregarded when the size of the dual space is large. On the other hand, the subgradient method initialized with a "good" first estimate allowed us to approximate the CHP. It should however be noted that the convergence rate of the subgradient method can possibly be an issue for practical implementation.
- Concerning the feasibility of the retrieved solution, for both the vanilla and the real case the SOC relaxation gives satisfaction since the solutions are physically meaningful.
- Concerning the uplift in the real Italian case, it is a nonconvex scenario with

a nonzero associated uplift. The approximate of the CHP gives the lowest one which is about 45% smaller than the highest uplift obtained with the restricted approach and about 35% smaller than the uplift computed with the dispatchable approach. In comparison with the total profits generated by the producers and the network operator, these uplifts are negligible.

5.2 Improvement paths

There are several interesting directions for future research:

- The analysis performed here takes only one country and a few time steps into account. Increasing the number of time steps and analyzing different countries and periods of the year would allow to obtain more information about "how bad" the required uplifts are. Specifically, comparing the welfare obtained with the pricing schemes presented here with the welfare obtained with `EUPHEMIA` would be of the greatest interest in order to assess their potential benefit.
- As highlighted, in this analysis, the SOC relaxation gives satisfaction since physically meaningful solutions were obtained which was not guaranteed. Extending the analysis would also allow to test this element and, if it happens that we simply had luck in our analysis, trying other relaxations could be considered.
- Since the CHP seems to be the most desirable pricing scheme, trying other ways to retrieve it would be interesting. In the same idea of what was done here, other dual function optimization techniques could be exploited but finding a way to formulate the primal approach efficiently could possibly be a better idea and would also allow to obtain the CHP more quickly.

Bibliography

[1] G. Leclercq O. Devolder Frederik G. A. Ashouri, P. Sels and R. Dhulst. Network and market models: preliminary report. *Technical report*, 2017.

[2] H. Gerard & al. Coordination between transmission and distribution system operators in the electricity sector: A conceptual framework. *ELSEVIER*, September 2017.

[3] R. O'Neil & al. Efficient market-clearing prices in markets with nonconvexities. *European Journal of Operational Research*, 2005.

[4] S. Boyd. course slide, convex optimization ii, stanford university, [online]:`https://web.stanford.edu/class/ee364b/lectures/subgrad_method_slides.pdf`.

[5] B. Cornélusse D. Ernst. How the european day-ahead electricity market works. March 2017.

[6] M. Farivar and S. Low. Branch flow model: Relaxations and convexification: Part i. *IEEE Transactions on Power Systems, 28(3)*, page 2554–2564, August 2013.

[7] European Distribution System Operators for Smart Grids. Flexibility: The role of dsos in tomorrow's electricity market. *Technical report*, 2014.

[8] L. Wolsey G. Nemhauser. Integer and combinatorial optimization. 1988.

[9] W. Hogan. Workshop on optimization and equilibrium in energy economics. pages 32–33, January 2016.

[10] PJM Interconnection. Proposed enhancements to energy price formation. *Technical report*, November 2017.

[11] R. Jabr. Radial distribution load flow using conic programming. *IEEE Transactions on Power Systems, 21(3)*, pages 1458–1459, August 2006.

[12] F. Wu M. Baran. Network reconfiguration in distribution systems for loss reduction and load balancing. *IEEE Transactions on Power Delivery*, April 1989.

[13] M. Van Vyve M. Madani. Computationally efficient mip formulation and algorithms for european day-ahead electricity market auctions. *European Journal of Operational Research*, 2015.

[14] I. Mezghani and A. Papavasiliou. A mixed integer second order cone program for transmission-distribution system co-optimization. *IEEE Milan PowerTech*, 2019.

[15] Y. Nesterov. Introductory lectures on convex optimization - a basic course. *Springer Science + Business Media, LLC*, 2004.

[16] W. Hogan P. Griki and S. Pope. Market-clearing electricity prices and energy uplift. December 2007.

[17] A. Papavasiliou. Optimization models in electricity markets [textbook course "linma2415", uclouvain].

[18] L.L.C PJM Interconnection. Update on pjm price formation efforts, [online]:`https://pjm.com/-/media/committees-groups/stakeholder-meetings/pjm-miso-joint-common/20190222/20190222-item-03-price-formation-and-cpf-update.ashx`. 2019.

[19] N. Stevens. Models and algorithms for pricing electricity in unit commitment. *Master's thesis*, 2016.

[20] J. Taylor. Convex optimization of power systems. *Cambridge University Press*, Septembre 2014.

[21] SmartNet's website. [online]: `http://smartnet-project.eu`.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain
Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/epl**