Erasmus University Rotterdam
Erasmus School of Economics
Master Thesis Econometrics and Management Science

Optimal Expansion Planning of an Integrated Energy System with Investment Options

Wessel Boosman (589273)

Supervisor:	Dr. O. (Olga) Kuryatnikova
	N. (Nicholas) Fuchs, MSc.
	N. (Natapon) Wanapinit, MSc.
Second assessor:	Dr. E.R. (Ruben) van Beesten
Date final version:	9th August 2024

Abstract

This work presents a budget constrained energy system expansion problem with integer investment options, where the objective is to minimize the operational cost related to electricity generation. This problem is modelled as a Mixed-Integer-Quadratic-Programming (MIQP) problem. The optimal objective value represents an optimal set of investments and their corresponding Optimal Energy Flow (OEF) schedule. Most solution approaches rely on repeatedly optimizing the OEF problem. Therefore, this thesis explores whether the distributed optimization algorithm, Alternating Direction Method of Multipliers (ADMM), when applied to the OEF problem, could increase the computational efficiency of the overall problem. It is found that, based on multiple benchmark instances, this approach could not outperform the commercial solver 'Gurobi'. On the other hand, the investment optimization (i.e., identifying the optimal investment strategy that minimizes the objective value) could not be solved in a reasonable time by a commercial solver. Consequently, two heuristics are presented and, based on a medium-sized energy model with *12* nodes and *8760* time-steps, the *Relax & Fit* heuristic achieved an objective value only 3.26% higher than the lower bound. Additionally, it is demonstrated that the selected investments corresponding to this result show a strong connection with the marginal costs defined at the nodes of the system.

Acknowledgement

This thesis marks the end of my graduate program, and I am grateful for the knowledge and skills I have gained during this time, as well as the doors it has opened. I also want to extend my thanks to the Fraunhofer Institute for supporting me with my master's thesis. Especially, I would like to thank Nicholas Fuchs and Natapon Wanapinit for their constant support and guidance. Furthermore, I would like to thank Olga Kuryatnikova for her constructive feedback and excellent guidance throughout the thesis.

Contents

1 Introduction

1.1 Context & Motivation

Electricity generation will be characterized by a shift away from a fossil fuel based generation in the coming decades. This offers a unique opportunity for policymakers to shape the future energy infrastructure. Distributed and unpredictable energy generation by renewable sources and technologies that can convert one energy-carrier into another (e.g. Power-to-Gas technologies, heat pumps and Combined Heat and Power plants) will reshape the dynamics of the energy market. The price paid for this next generation infrastructure will be in terms of complexity, generation will become less predictable, load profiles will need to adapt to supply, and heat-pumps put additional stress on the electricity network. Therefore, with the current state of distributed energy generation from renewable sources, it becomes increasingly important to model the integrated energy system (Krause, Andersson, Fröhlich & Vaccaro, 2011). This work is in collaboration with the *'Fraunhofer-Institut für Solare Energiesysteme'* (ISE) department of the *Fraunhofer-Gesellschaft* research institute. This department is, among other activities, concerned with energy optimisation problems. To be more specific, they are focusing on the most cost-effective way of satisfying certain future energy goals. These objectives could be related to expectations of energy and heat demand or limiting greenhouse gas (GHG) emission, for example. Due to the complexity of their energy models, these optimizations are computationally challenging, and consequently, for large projects the models are to a certain degree limited in terms of input size, spatial resolution, temporal resolution, or all three. Even with these limitations, the models are still too large and resource-intensive to solve efficiently. This thesis aims at addressing this issue by modelling an integrated energy model (i.e. a model with more than one energy-carrier), and presenting an effective algorithm that can determine the optimal allocation of financial resources such that public utility is optimized. This utility function can either be expressed in monetary value or the reduction of GHG's.

1.2 Research direction

This work focuses on the optimal investment problem in the field of energy infrastructure limited by financial resources and/or emission constraints. The infrastructure considered consists of an integrated system with an electrical and a thermal network. The difficulty of this problem lies in the fact that the utility of candidate investments is an unknown function that depends not only on the system in which it is potentially applied, but also on other planned investment. For example, the utility of renewable energy can be considerably higher when installed together

with storage facilities. Secondly, increased grid capacity can increase usage of 'cheap' renewable energy. Therefore, this problem is a combinatorial problem in which the utility is associated with the combination of investments rather than its individual components. Understandably, the numbers of possible combinations grows exponentially with the number of investment options. For this reason, we expect that a commercial solver will not be able to find the optimal set of investments in a reasonable time. Moreover, to find the utility (i.e. the positive effect a combination of investments has on the system measured either in monetary value or value related to avoiding GHG emissions) we need to determine its effect against a benchmark model in which we should consider at least a full one-year time horizon to capture seasonal effects. This can be accomplished by running a full-year Optimal Energy Flow (OEF) problem. Overall, we consider a large Mixed Integer Programming (MIP) problem aimed at maximizing utility. Conceptually, this problem can be divided into two parts: an integer investment problem aimed at finding the optimal investment strategy and secondly, a corresponding full-year OEF problem for finding the optimal energy flow that supports the objective. We expect that the OEF problem needs to be evaluated often and even small computational efficiency gains could result in a faster algorithm. Therefore, we explore the possibility of decomposing this problem into smaller sub-problems.

In summary, the research question and related sub-questions are:

How can an integer investment problem in the context of energy optimization be solved efficiently?

- What methods would be most suitable to decompose the OEF problem and how do they perform in comparison to an out-of-the-box solver?
- What heuristic or relaxations can be proposed and efficiently used for the investment problem?

1.3 Related work

The scheduling of electricity is a widely researched problem due to its importance in daily electricity management, and is known as the Optimal Power Flow (OPF) problem. In reality, energy transport, whether in the form of electricity, natural gas or heat, is subjected to specific physical behavior which can seldom be captured by convex formulations. For example, gas flow is realistically non-convex due to its pressure losses (Xu et al., 2020) along the pipe. Electricity transmission realistically involves reactive power, transmission losses and power dependency on the voltage squared (Chatzivasileiadis, 2018), making the flow equation non-linear and non-convex. In this thesis however, the focus is not necessarily on the exact power scheduling with high temporal resolution, but rather on a simplification detailed enough to evaluate the effect of investments. This can, for example, be done by linearizing the quadratic flow equations, which results in a simplified convex model known as the DC Power Flow Model (Chatzivasileiadis, 2018). Moreover, Moehle, Shen, Luo and Boyd (2019) propose a method where not only the flow equations but en entire energy system can be modelled by a system of individual components

that are all described by a convex objective function and convex constraints. These individual components, also called devices, represent participants in the energy system and include, for example, generators, batteries, and transmission lines. They then demonstrate that this model can be used to determine the optimal power flow in deterministic cases as well as under uncertainty. The advantage of modeling an energy system based on a network of devices is that the model can easily be adjusted by adding or removing them. This makes it a very suitable method for evaluating various investment strategies. Moreover, Kraning (2014) demonstrates that its structure allows for distributed optimization using the ADMM algorithm. For these reasons, this modelling approach will be used to construct an energy system in this thesis. However, it will be extended by integrating a thermal network and devices that couple the electrical grid with the thermal network.

Research with respect to optimal investment in an energy system is, for example, done by Dvorkin, Kazempour, Baringo and Pinson (2018). To find the optimal capacity expansion plan for certain electrical generators, the author introduces a two-stage optimization problem consisting of short- and long-term decisions that are both subjected to uncertainty (e.g. operational risk and growth demand risks). The problem is, therefore, a multi-stage stochastic optimization problem. It requires agreement on the shared first-stage decision variables, which determine the capacity to be installed for three types of electrical generators: coal, gas, and wind. The objective in this paper is somewhat similar to the problem we are facing; but it only considers three investment options and does not consider integer variables (i.e. any fractional capacity can be installed). Instead, its focus is more on how to incorporate uncertainty, which will not be considered in this thesis. Alternatively, in the work of Zhang, Shahidehpour, Alabdulwahab and Abusorrah (2015), the authors construct an optimal expansion planning model that satisfies electricity, heat and natural gas demands over a fixed time horizon and show that optimizing with respect to these commodities simultaneously offers better results compared to planning them individually. However, they only consider continuous capacity expansion and use time-steps of one month, which doesn't capture the fluctuations of renewable energy generation.

1.4 Outline

The remaining of this work is be structured as follows, the next chapter describes distributed optimization methods. The focus here is on the 'Alternating Direction Method of Multipliers' (ADMM) and especially the 'Exchange ADMM' variant as this is a recurring algorithm in the remainder of the thesis. Subsequently, in Chapter 3 the OEF problem is introduced, how it can be decomposed, and the distributed optimization algorithm considered for solving it. Secondly, the optimization problem of selecting investments that maximize public utility is introduced. Additionally, two heuristics are provided for solving this problem. Lastly, an example is introduced on which the proposed methods are tested. Then, in Chapter 4 we will present the results. Subsequently, Chapter 5 will conclude the work and provide recommendations for the ISE department of the *Fraunhofer-Gesellschaft*. Additionally, Chapter 5 will discuss several limitations of the methods and assumptions, and propose three ideas for future research.

2 | Distributed Optimization

This chapter provides an overview of a few algorithms developed to decompose and distribute large optimization problems. These algorithms can be used to address the first research question of this work. These algorithms will be divided into two categories, gradient based and cutting planes based algorithms. The two categories share the same goal of solving large optimization problems efficiently but their working principles differ. Although the focus will be on gradient-based algorithms, a comprehensive comparison will be made to explain the differences.

2.1 Gradient based algorithms

Within distributed optimization, gradient based algorithms are powerful when the objective function can be decomposed. This approach enables parallel computing, in which sub-problems are independently solved, and uses gradient steps to move towards the optimal solution. One of these algorithms is the 'Alternating Direction of Multipliers' (ADMM), which will be discussed in the next subsections.

Alternating Direction Method of Multipliers

In the ADMM algorithm the objective function is decomposed among the primal variables. Let us consider the following problem, where f and g are convex functions and X and Z are convex sets:

$$\begin{aligned} \text{minimize} \quad & f(x) + g(z) \\ \text{subject to} \quad & Ax + Bz = c \\ & x \in X \\ & z \in Z \end{aligned} \tag{2.1}$$

The decision variables are connected through the linking equality constraints. Although the variables are connected, the main idea behind the ADMM algorithm is that initially, it does not strictly enforce whether the linking constraints hold or not. Instead, the sub-problems with respect to x and z are solved independently to optimality; and in these sub-problems, any violation of the linking constraints is quadratically penalized. Therefore, the formulation is different from the conventional Lagrangian relaxation and is referred to as the augmented Lagrangian, as shown in Equation (2.2). The quadratic penalty term gives rise to very favourable convergence properties, as described in earlier work of Boyd (2010). Also note that the quadratic L_2 regularization of a convex function (or constraint) is still convex because the second derivative is always non-negative, and thus, the augmented Lagrangian function is also convex.

$$L_\rho(x, z, \lambda) = f(x) + g(z) + \lambda^T(Ax + Bz - c) + \frac{\rho}{2}\|Ax + Bz - c\|_2^2 \tag{2.2}$$

Note. Often it is easier to scale the dual variable of Equation (2.2) by $^1/_\rho$ and include the dual penalty in the quadratic regularization:

$$L_\rho(x, z, \lambda) = f(x) + g(z) + \frac{\rho}{2}\|Ax + Bz - c + \left(\frac{1}{\rho}\lambda\right)\|_2^2 \tag{2.3}$$

This formulation is equivalent (but not equal) to Equation (2.2). At first sight, this may not be immediately evident, and the derivation is often not explicitly provided. Therefore, for completion, we derive here the scaled form of ADMM, from the standard formulation and show that the two formulations are indeed equivalent.

Consider the Lagrangian relaxation part of Equation (2.2):

$$\lambda^T r + \frac{\rho}{2}\|r\|_2^2 \tag{A}$$

Where r is the primal residual $Ax + Bz - c$. We will show that this is equivalent to:

$$\frac{\rho}{2}\|r + \lambda_{scaled}\|_2^2 \tag{B}$$

Where $\lambda_{scaled} = \frac{\lambda}{\rho}$.

Including the linear term of Equation (A) in the quadratic term gives:

$$\frac{\rho}{2}\|r - \frac{\lambda}{\rho}\|_2^2 - \frac{\rho}{2} \cdot \frac{\lambda^2}{\rho^2}$$

Substitute λ_{scaled} in this formulation gives us:

$$\frac{\rho}{2}\|r - \lambda_{scaled}\|_2^2 - \frac{\rho}{2}\lambda_{scaled}^2$$

Which is in the optimization steps of ADMM equivalent to:

$$\frac{\rho}{2}\|r - \lambda_{scaled}\|_2^2 - \text{constant}(\lambda_{scaled})$$

The constant does not play a role in the optimization with respect to x and z and therefore scaling the dual variable leads to equal results.

Finding the optimal dual variable can be achieved by applying the dual ascent method, where a gradient step is made each iteration based on the violation of the relaxed constraint. Applying dual ascent on the augmented Lagrangian instead of the standard Lagrangian is referred to as ADMM, and its algorithm is summarized in Algorithm 1. This method will converge to the global optimum because the dual ascent method pulls the dual variable toward its optimal value, while the primal variable optimization does this for the primal variables. In addition, strong duality means that there is no duality gap, and the optimal primal solution correspond directly

to the dual solution.

Also important to note is that in Algorithm 1 the x and z update can be independently computed, which means the algorithm can take advantage of parallel computing. More information on the convergence properties and the optimality conditions of ADMM can be found in (Boyd, 2010).

Algorithm 1 ADMM algorithm

Initialize $L_\rho(x, z, \lambda) = f(x) + g(z) + \lambda^T(Ax + Bz - c) + \frac{\rho}{2}\|Ax + Bz - c\|_2^2$
Initialize $x^0, z^0, \lambda^0 \leftarrow x^{init}, z^{init}, 0$
$k \leftarrow 0$
$\mu \leftarrow \epsilon + 1$ ▷ Set μ to an arbitrary number above ϵ for the first iteration
while $\mu \geq \epsilon$ **do**
    $x^{k+1} \leftarrow \text{argmin}_{x \in X}\, L_\rho(x^k, z^k, \lambda^k)$
    $z^{k+1} \leftarrow \text{argmin}_{z \in Z}\, L_\rho(x^{k+1}, z^k, \lambda^k)$
    $\lambda^{k+1} \leftarrow \lambda^k + \rho(Ax^{k+1} + Bz^{k+1} - c)$
    $\mu \leftarrow \|\lambda^{k+1} - \lambda^k\|$
    $k \leftarrow k + 1$
end while

Consensus ADMM

In many real-world applications the objective function can be formulated as the sum of multiple sub-problems, in which the z-variable reflects a decision that has to be agreed upon among the sub-problems. These kinds of problems arise, for example, in two stage stochastic optimization. You could decompose this optimization problem by scenario, in which all sub-problems need to agree on this first-stage decision. In other words, there needs to be consensus among the sub-problems regarding one or more of the variables. When using ADMM for these kind of problems, the algorithm is also knows as 'Consensus ADMM' (Boyd, 2010). In this section a closer look is taken into this special case, showing that when the z-variable is not part of the objective function, the z-update can be solved analytically. Subsequently, the 'Exchange ADMM' algorithm will be derived from 'Consensus ADMM'. Consider the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{i \in \mathcal{N}} f_i(x_i) \\ \text{subject to} \quad & h_i(x_i) \leq C_i, \qquad \forall i \in \mathcal{N} \\ & x_i - z = 0, \qquad \forall i \in \mathcal{N} \end{aligned} \tag{2.4}$$

It can be recognized that the first set of constraints are local constraints to the sub-problem, and the second set of constraints are considered the (difficult) coupling constraints. The augmented Lagrangian with respect to the coupling constraint can be formulated as follows:

$$L_\rho(x, z, \lambda) = \sum_{i \in \mathcal{N}} f(x_i) + \lambda^T(x_i - z) + \frac{\rho}{2}\|x_i - z\|_2^2 \tag{2.5}$$

The subproblem with respect to the x_i-update is the minimization of the augmented Lagrangian with respect to x_i subjected to the local constraints. The z-update and the dual variable update are very similar to the steps from Algorithm 1, and are given below:

$$\left\{\begin{aligned} x_i^{k+1} &= \min_{x_i}\left(f_i(x_i) + (\lambda_i^k)^T x_i + \frac{\rho}{2}\|x_i - z^k\|_2^2\right) \\ \text{s.t.} \quad & h_i(x_i) \leq A_i \end{aligned}\right\}, \qquad \forall i \in \mathcal{N} \tag{2.6a}$$

$$z^{k+1} = \min_z \left(-\sum_{i\in\mathcal{N}} (\lambda_i^k)^T z + \frac{\rho}{2}\sum_{i\in\mathcal{N}} \|x_i^{k+1} - z\|_2^2\right) \tag{2.6b}$$

$$\lambda_i^{k+1} = \lambda_i^k + \rho(x_i^{k+1} - z^{k+1}) \tag{2.6c}$$

Again, the x-update can be fully parallelized. Furthermore, the z-update can be simplified by finding an analytical solution to the minimization problem. Since the optimization problem is convex, quadratic and unconstrained, there is only one optimal solution, which corresponds to its unique zero gradient point. Taking the derivative to z and setting the expression to zero results in the optimal z-update. The result is given below. For completeness, the derivation is included subsequently, as it is usually not explicitly provided.

$$z^{k+1} = \min_z \left(-\sum_{i\in\mathcal{N}} (\lambda_i^k)^T z + \frac{\rho}{2}\sum_{i\in\mathcal{N}} \|x_i^{k+1} - z\|_2^2\right) = \frac{1}{\rho}\overline{\lambda}^k + \overline{x}^{k+1} \tag{2.7}$$

Derivation of Equation (2.7)
Here it is demonstrated that Equation (2.7) is the optimal z-update in the consensus ADMM algorithm by showing that it corresponds to the zero-gradient point. Since optimization Problem (2.6b) is unconstrained and convex quadratic, the gradient with respect to z of Equation (2.5) needs to be equal to zero:

$$\begin{aligned} \nabla_{z^*}\left(-\sum_{i\in\mathcal{N}} (\lambda_i^k)^T z^* + \frac{\rho}{2}\sum_{i\in\mathcal{N}} \|x_i^{k+1} - z^*\|_2^2\right) &= 0 \\ -\sum_{i\in\mathcal{N}} \lambda_i^k - \rho\sum_{i\in\mathcal{N}} (x_i^{k+1} - z^*) &= 0 \\ \sum_{i\in\mathcal{N}} \lambda_i^k + \rho\sum_{i\in\mathcal{N}} x_i^{k+1} &= \rho N z^* \\ \frac{1}{\rho N}\sum_{i\in\mathcal{N}} \lambda_i^k + \frac{1}{N}\sum_{i\in\mathcal{N}} x_i^{k+1} &= z^* \\ \frac{1}{\rho}\overline{\lambda}^k + \overline{x}^{k+1} &= z^* \end{aligned} \tag{2.8a}$$

From here it can be seen that Equation (2.7) is the optimal solution for the update of the global variable z.

Despite the nice closed form for the z-update, the consensus ADMM algorithm can be

simplified further. Averaging Equation (2.6c) over $i \in \mathcal{N}$ results in:

$$\overline{\lambda}^{k+1} = \overline{\lambda}^{k} + \rho(\overline{x}^{k+1} - z^{k+1})$$

Substituting the closed form of z^{k+1} from Equation (2.7), shows that $\overline{\lambda}^{k+1} = 0$. In other words, the average of the dual variable update is zero after the first iteration. If the initial values of the dual variables are zero, the initial average of the dual variables are zero ($\overline{\lambda}^{k} = 0$) and consequently, $z^{k+1} = \overline{x}^{k+1}$. After the first iteration, this also implies that $z^{k} = \overline{x}^{k}$. Now, it can be easily seen that the consensus ADMM algorithm can be described in two instead of three updates:

$$\left\{ \begin{array}{l} x_i^{k+1} = \min\limits_{x_i} \left(f_i(x_i) + (\lambda_i^k)^T x_i + \dfrac{\rho}{2} \|x_i - \overline{x}^k\|_2^2 \right) \\ \qquad \text{s.t.} \quad h_i(x_i) \leq A_i \end{array} \right\}, \qquad \forall i \in \mathcal{N} \tag{2.9a}$$

$$\lambda_i^{k+1} = \lambda_i^k + \rho(x_i^{k+1} - \overline{x}^{k+1}) \tag{2.9b}$$

Although each sub-problem is free to decide on its x_i variable, they are penalized based on their distance to the average value (i.e. averaged over all participants) of the global variable, which is determined from the results of all sub-problems in the previous iteration. In other words, 'participant A' might minimize its objective by taking on a x_i value that deviates significantly from the average in the previous iteration, thereby influencing the average value of the global variable directly. The next iteration, all other 'participants' in the system will be penalized for their distance to the average value, and are therefore indirectly respecting the optimal choice of 'participant A'. This continues until no participant of the system can deviate from the common value such that its personal gain outweighs the negative impact on the overall minimization. This last sentence shows a nice interpretation of the overall algorithm by looking closely at the penalty term in the augmented Lagrangian.

Although, this result is achieved by reasoning, it can actually be proved that the optimal solution satisfies a 'Nash equilibrium' point, i.e. a solution in which none of the participants of a system can gain from unilaterally deviating from the equilibrium point (Holt & Roth, 2004). In the next section the ADMM variant 'Exchange ADMM' is derived from the result in Equation (2.6). An explanation will be provided as to why this variant of the algorithm can be used in the area of energy optimization and how the optimal solution corresponds to a Nash equilibrium.

Exchange ADMM

Suppose an optimization problem involves participants who either buy or sell a certain product. When all participants aim to minimize their respective cost functions, which can be related to utility or profits depending on the participant, the formulation of this optimization problem can be described as follows:

$$\begin{aligned} \text{minimize} \quad & \sum_{i \in N} f_i(x_i) \\ \text{subject to} \quad & h_i(x_i) \leq C_i \qquad \forall i \in N \\ & \frac{1}{|N|} \sum_{i \in N} x_i = 0 \end{aligned} \tag{2.10}$$

This problem naturally occurs in (financial) markets, where the objective function can be interpreted as the minimization of the aggregated cost/utility function of all the participants in the market (i.e. buyers and sellers). These cost functions represent the amount a certain participant is willing to sell/buy and for which price. The internal constraints could represent the number of financial instruments or the amount of commodity a participant is willing to trade. The third constraint can be interpreted as the balancing constraint, forcing that whatever is bought at the market is also sold, which means that also the average of the exchange should equal zero. Due to its application, literature describes this problem also as the 'Market Clearing' problem or as 'Walrasian Auction' (Joyce, 1984). How this problem connects to energy models will be the focus of the subsequent chapter. Here, it will be explained how these problems could benefit from a distributed optimization algorithm.

Looking carefully to Problems (2.4) and (2.10), it is evident that the latter formulation is a variant of Formulation (2.4) in which:

$$z = x_i - \frac{1}{\mathcal{N}} \sum_{i=0}^{\mathcal{N}} x_i = x_i - \overline{x} \tag{2.11}$$

Substituting this result in Equation (2.9a) and (2.9b), results in the final form of the 'Exchange ADMM' algorithm, which in the scaled form can be summarized as in Equation (2.12) below. The reason of using the scaled form instead of the standard dual variable is that it reduces the information exchange between the coordinator and the sub-problems. Instead of handing the average value ($\overline{x}^{k-1}$) **and** the dual variable (λ^{k-1}) back to the sub-problems, in the scaled form only the locally computed (by the coordinator) penalty term $x^{k-1} - \overline{x}^{k-1} - \lambda^{k-1}$ needs to be returned. This makes implementation of the algorithm easier and more efficient.

$$\left\{ \begin{aligned} x_i^{k+1} = \min_{x_i} & \left(f_i(x_i) + \frac{\rho}{2} \| x_i - (x_i^k - \overline{x}^k - \lambda_i^k) \|_2^2 \right) \\ \text{s.t.} \quad & h_i(x_i) \leq A_i \end{aligned} \right\}, \qquad \forall i \in \mathcal{N} \tag{2.12a}$$

$$\lambda_i^{k+1} = \lambda_i^k + \rho \overline{x}^{k+1} \tag{2.12b}$$

Instead of decomposing Problem (2.10) into $\mathcal{N}$ sub-problems and solving it using the algorithm above, there is a second approach. In work of (Kazempour, 2024), it is argued that the Market Clearing problem can also be written as a Nash equilibrium problem. This formulation consists of $\mathcal{N} + 1$ optimization problems, one for each participant plus a 'Price-setter' optimization. Then, Kazempour (2024) shows that the system of equations derived from the KKT conditions for each of those optimization problems is equal to the KKT conditions of Problem

(2.10), indicating that both perspectives of the Market Clearing Problem are equivalent and that the optimal solution of Problem (2.10) satisfies a Nash equilibrium. Chapter 3 demonstrates how this result connects to duality theory and its application in energy modeling.

Penalty term

The choice of the penalty term ρ depends on the scale of the data, but should nevertheless be a positive non-zero number. There is no fixed rule of thumb for estimating the right parameter. Instead, ρ could either be set to 1, and the data should be scaled such that a reasonable convergence rate is achieved, or ρ can be determined empirically through parameter tuning.

2.2 Cutting-plane algorithms

In contrast to gradient based algorithms, under which ADMM falls, cutting-plane based algorithms are not making a gradient-based step in each iteration to find an optimal solution. Instead, constraints or variables that are important to include for finding the optimal solution are added iteratively. For example, in Benders decomposition, the problem is separated in a master problem containing the (difficult) linking variables, and sub-problems that optimize for the remaining variables in which the linking variables are thought of as fixed. During each iteration, the dual variables corresponding to the linking constraints in the sub-problems are communicated to the master problem. The master problem then uses these dual variables to create a new cut (constraint) that approximates the original objective function as closely as possible. This method is especially useful for MILP problems as long as the integer variables are restricted to the master problem (Benders, 1962). These situations occur frequently in two-stage stochastic optimization. Based on two-stage optimization, its also clear to see how the gradient-based consensus ADMM differs form Benders decomposition. The former decomposes the primal variables and relaxes, but penalizes any violation of the 'linking constraints'. Each sub-problem has the freedom to decide on their primal variables as long as it positively impacts the common objective. In Benders decomposition, on the other hand, the 'difficult' part of the problem is fixed in each of the sub-problems. The role of the sub-problems is to generate new constraints to the master problem that will improve the objective function.
Alternatively, Dantzig-Wolfe decomposition is a cutting-plane algorithm that is equivalent to Benders decomposition when considering the dual problem. Instead of constraints, it generates columns (i.e. variables) to include in the master problem. In the context of power flow optimization, these columns could reflect a (partial) power flow that should be beneficial in the original problem. Additionally, the formulation of a power flow problem can be described by a block angular restriction matrix (necessary structure for Dantzig-Wolfe), in which each block describes the set of constraints associated with an electrical bus (i.e., a node where multiple electrical connections intersect). The transmission lines connecting electrical buses generate a set of linking constraints (Bastianel, Ergun & Hertem, 2022). Therefore Dantzig-Wolfe could be a suitable alternative for our problem.

2.3 Algorithm choice

Whether to use a gradient-based algorithm in the form of ADMM or a cutting-plane method depends on the problem. Our problem consists of an energy optimization problem with investment opportunities. Here, the integer investment problem can not be seen separately form the operational, or OEF problem. Although the OEF problem can be modelled as a convex optimization problem, it can still become quite large. Assuming this problem needs to be evaluated frequently, solving it efficiently could bring significant time gains. Benders decomposition will not be usable in this context because it requires full separation of the sub-problems while the transmission lines generate a lot of linking constraints between sub-problems. Both Dantzig-Wolfe decomposition and ADMM are viable options. However, the ADMM algorithm is preferred. While Dantzig-Wolfe could decompose the system per energy hub or bus, Kraning (2014) has shown that an ADMM algorithm enables to go one step further, and decomposes the problem by each component in the system, such as a wind turbine, a transmission line, batteries or thermal and electrical loads. Furthermore, studies by Bastianel et al. (2022) indicate that for small to medium sized problems a Dantzig-Wolfe algorithm for solving the DC OPF problem is significantly slower compared to the commercial solver '*Gurobi*'.
Therefore, this thesis builds upon work earlier done by Kraning (2014), expanding it by integrating a thermal network. The decomposition and implementation of this algorithm will be the subject of the next chapter.

3 Energy system optimization with investment options

The central theme in this thesis is the co-optimization of operation and investments within an energy model. This chapter will introduce the problem formulation, present multiple solution approaches, and provide an example on which these solution methods will be tested. Before formulating the optimization problem, Section 3.1 first describes the adopted method of Moehle, Busseti, Boyd and Wytock (2019) on modelling an energy system based on devices, extended by introducing the modeling description of 'Combined Heat and Power' (CHP) plants. Subsequently, Section 3.2 formally introduces the optimization problem. This optimization consists of selecting the optimal set of investments, such that it minimises the overall cost of energy generation while satisfying demand. This section will also show that this problem can not be seen independently from the OEF problem. Therefore, Section 3.3 will dive deeper into the OEF problem and show that an energy model based on the devices described in Section 3.1 can be decomposed and solved using the *'exchange ADMM'* algorithm. Thereafter, Section 3.4 addresses the investment optimization and introduces two heuristics for solving this non-convex problem. Lastly, Section 3.5 proposes an example model to evaluate the performance of the solution approaches.

3.1 Device decomposition

This section describes how an energy system can be modelled based on devices and nodes. The term *'Devices'* serves as an umbrella term for all participants in the model, and include, for example, generators, electrical loads, batteries, transmission lines, etc. Similarly, '*nodes*' denote connection points in the model where these devices are linked. This approach was earlier proposed in Moehle, Busseti et al. (2019), where the authors not only introduce the method but also define specific devices in terms of constraints and cost functions, some of which are adopted in this work. This method of modelling an energy system has the advantage that devices can easily be added or removed to change or expand the model. Moreover, entire networks interacting with the electricity grid can be integrated into the system using this approach. Examples include district heating networks, low- and high-voltage electricity grids, and devices producing hydrogen. This work considers an energy model consisting of an electrical network and a thermal network that are interconnected through CHP plants. A schematic figure of such a model is given in Figure 3.1.

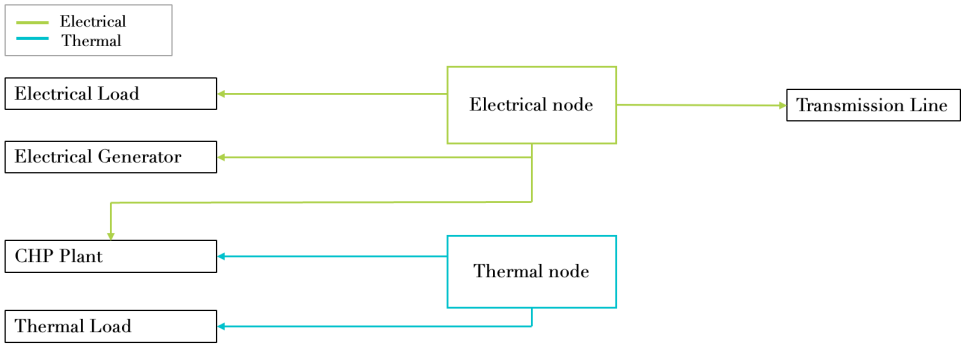

Figure 3.1: *Schematic view of modelling an energy network based on devices and multiple energy carriers*

3.1.1 Generators

In this section the generating devices are described based on their cost-function and constraints.

Conventional generators

Conventional generators are defined as generators that produce power from natural gas or other fossil fuels. In work from Durvasulu and Hansen (2018), the authors explain that the cost-function of these generators can be approximated by a quadratic function of the electrical output. This is because these generators consume more fuel per unit of electrical output when they operate away from their optimal operating point. This behaviour is described by the Heat-Rate Curve, which plots the amount of fuel required per MWh electrical output as a function of the total electrical output of the generator. The cost-function is repeated in Equation (3.1), in which p is the electrical output. The parameter β corresponds to the linear part of the cost-function and is associated with fuel cost. The parameter α represents the curvature of the cost function and describes how the cost per MWh changes with the operating level. Fixed costs are represented by γ.

$$f(p) = \alpha \cdot p^2 + \beta \cdot p + \gamma \tag{3.1}$$

Internal constraints for the generator include ramping constraints, as previously described by (Moehle, Busseti et al., 2019), as well as maximum capacity constraints. Ramping constraints limit how much a generator can change its output from one hour to the next, and these limitations usually vary depending on the generator's size and type. These internal constraints are summarized below:

$$\left.\begin{aligned} p^t - p^{t-1} &\leq R_{increase} \\ p^t - p^{t-1} &\geq R_{reduce} \end{aligned}\right\} \quad \rightarrow \textit{Ramping constraints for increasing and reducing operation respectively}$$

$$p \leq p_{max} \quad \rightarrow \textit{Maximum capacity constraint} \tag{3.2}$$

Boiler

A boiler is a device that can provide heat to a thermal node by burning fuel. In contrast to a conventional generator, a linear cost-function is assumed. While electrical output adheres to optimal operating points due to design choices regarding turbines and other mechanical equipment, heat production is a far more straightforward process of simply burning fuel. In other words, it is assumed that burning twice as much fuel results in twice as much thermal energy. The cost-function can therefore be described as follows:

$$f(q) = \beta \cdot q + \gamma \tag{3.3}$$

Furthermore, the operation of boilers is also subjected to ramping and capacity constraints.

Renewable generators

Renewable generators are devices that provide their maximum available capacity with zero cost associated with their generation. They are therefore described by only one constraint.

$$p = p_{available} \tag{3.4}$$

Electricity generation is calculated by multiplying the total installed capacity with the availability of renewable source, either sun or wind. The availability of renewable sources is determined based on data from Renewable Ninjas (Pfenninger & Staffell, 2016). In case of overproduction, the power dissipation device will dissipate the excess power to ensure feasibility of the model.

Combined Heat- and Power Plant

Combined Heat- and Power plants are fossil fuel based generators that connects electrical and thermal networks. They have increased efficiency compared to their non-CHP peers and are considered vital to bridge the coming decades to a fully net-zero energy infrastructure. Increased efficiency is not due to higher thermodynamic efficiencies, but rather, as the name suggest, by using excess heat from electricity generation to satisfy thermal demands. Considering excess heat as 'useful' increases efficiency as can be seen by the following efficiency equation:

$$\eta_{CHP} = \frac{W_e + Q_{TH}}{Q_{Fuel}}$$

Here, W_e represents the useful electrical work done by the generator, Q_{TH} represents the useful heat (i.e. thermal energy) provided by the plant, and Q_{Fuel} is the input heat. Modern CHP plants can achieve efficiencies (η_{CHP}) of approximately 80%, and are also able to move operation towards more power orientated or heat orientated, depending on demand. We model the CHP as an aggregated device consisting of a conventional power plant and a boiler. Here, the conventional power plant has an additional connection to a thermal node on which (free) excess heat is transported.

The amount of thermal energy that can be recovered from the generator is approximately equal to its electrical output. We conclude that based on the fact that modern electrical generators usually reach efficiencies up to approximately 40% (W_e/Q_{fuel}), while modern CHP plants can operate at 80% efficiency (based on the assumption there is demand for the excess heat). The purpose of the boiler is to generate heat when electricity demand is low but heat demand is high, which coincides with the facts CHP plants can shift their operation according to demand. Or in other words, when electricity is produced it is always accompanied by the generation of free excess heat, while heat can be produced without the generation of electricity. Therefore, the cost function of CHP plants is defined as follows:

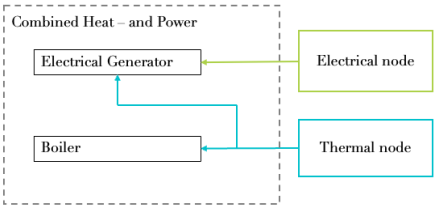

Figure 3.2: *A Combined Heat- and Power plant is modelled as an aggregated device consisting of a boiler and an electrical generator*

$$f(p, q) = \alpha_p \cdot p^2 + \beta_p \cdot p + \beta_q \cdot q + \gamma \tag{3.5}$$

In Equation (3.5), p denotes the electrical output of the system and q represents the thermal energy produced by the boiler. It's important to note that q does not necessarily represent the thermal energy exiting the plant, as it is complemented by excess thermal energy generated from electricity production. Furthermore, ramping constraints are defined for both internal devices, along with a combined maximum output constraint. These constraint are described earlier in this section.

3.1.2 Load Profiles

This section describes the electrical and thermal demand profiles.

Electrical Demand

Electrical loads are devices with a fixed hourly load profiles, a zero-cost function, and only one constraint:

$$p = p_{load} \tag{3.6}$$

This work considers three groups of electrical load profiles corresponding to households, commercial activity, and manufacturing. The differences between these profiles are summarized below and are visualized in Figure 3.3.

Residential & Commercial Residential and commercial load profiles are determined based on earlier work of Hellwig (2003), where the authors studied various types of load profiles, and summarized their work into a Python library that facilitates the construction of annual load profiles. The load profile for commercial activity is composed of a weighted average of different commercial sectors, such as office buildings, restaurants, bakeries, etc. An overview on the exact build-up can be found in Hellwig (2003).

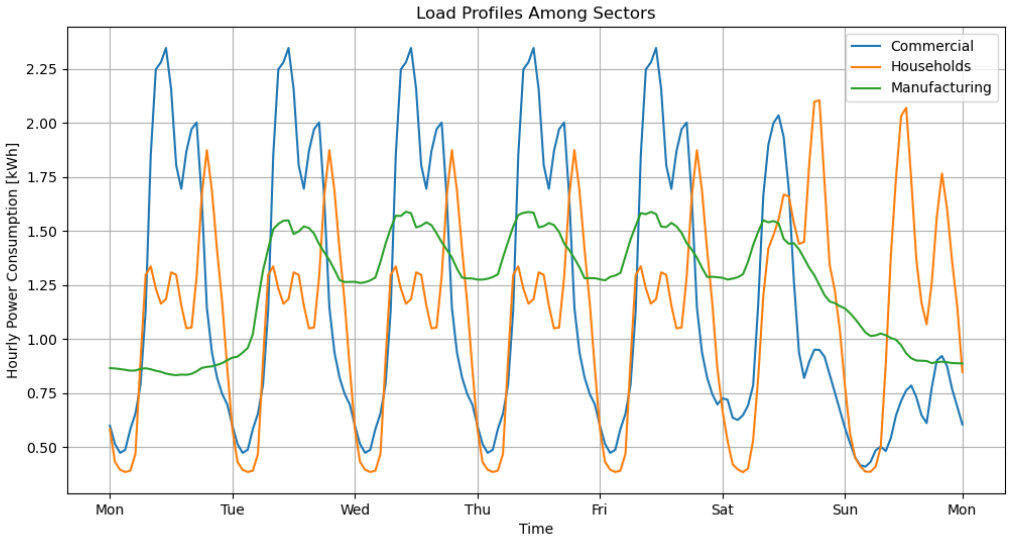

Figure 3.3: *For illustrative purposes, hourly power consumption [kWh] for the Commercial sector, Households, and the Manufacturing sector for a typical week. all scaled to have an annual demand of 10MWh.*

Manufacturing The load profile for the manufacturing sector is constructed as a weighted average of multiple manufacturing industries. Data used to construct this load profile are provided by Bellinguer, Girard, Bocquet and Chevalier (2023). The annual hourly dataset is normalized and can be scaled to match any desired annual demand.

Thermal Demand

Thermal loads represent fixed hourly load profiles that simulate energy usage by households for heating purposes. These load profiles are generated based on weather data from 2019 and are sourced from Pfenninger and Staffell (2016). Whether the thermal demand is satisfied by district heating or a heat-pump (thermal demand satisfied by domestic gas boilers is neglected as the natural-gas network is not modelled) affects the total energy use due to differences in efficiencies. Heat-pumps convert electrical energy into heat and are therefore connected to the electrical grid. Modern heat-pumps have a Coefficient of Performance (COP) of above three, meaning that every joule electrical energy can add three joules of heat to an environment. Thermal demands satisfied by district heating use (excess) heat from CHP plants and are therefore connected to a thermal network. Although heating an environment using district heating requires more energy than heat-pumps, it can still be extremely efficient due to the use of heat that would otherwise be wasted. Efficiency assumptions are summarized in Table 3.1.

Table 3.1: *Assumed efficiencies for satisfying thermal demand.*

Parameter	Value
Energy loss District Heating	0.1
Efficiency District Heating	0.9
$COP_{heating}$ Heat-Pump	3
$COP_{cooling}$ Heat-Pump	2

3.1.3 Transmission lines & Heat pipes

Transmission lines are connection devices between two nodes. Each transmission line has two input variables: the transported power from and to both connected nodes. Constraints include that the line operates within its limits and that the power subtracted by one node equals the power provided by the other. In reality, power losses occur and are dependent on the reactive power, which in turn depends on the current squared (Chatzivasileiadis, 2018). Transferred power is proportional to current, and thus it is fair to assume that transmission losses depend on power squared when voltage is fixed. As mentioned earlier, these flow equations are non-convex, but are often linearized, resulting in the DC Power Flow Model. To simulate power losses in a convex model, a cost-function is used that depends on the transferred power squared. This penalizes large power flows and simulate the additional cost made associated with these losses. This was earlier proposed in work of Moehle, Shen et al. (2019). The cost-function of transmission lines is defined in Equation (3.7). Here, α determines the cost associated with power losses. In reality, this parameter will depend on various factors such as cable diameter, voltage, length, and materials.

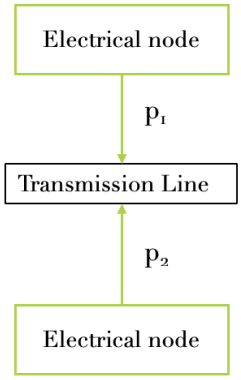

Figure 3.4: *Schematic depiction of a transmission line.*

$$f(p) = \alpha \cdot p^2 \tag{3.7}$$

The constraints of the transmission line are summarized below:

$$\left.\begin{aligned} \frac{p_1 - p_2}{2} &\leq p_{max} \\ \frac{p_2 - p_1}{2} &\leq p_{max} \end{aligned}\right\} \quad \rightarrow \textit{Maximum capacity constraints}$$
$$p_1 + p_2 = 0 \quad \rightarrow \textit{No energy loss} \tag{3.8}$$

Because transmission lines can transport energy among nodes, they can be viewed as devices that equalize energy prices across the network.

Transport of heat by water (or steam) could be be modelled by the same sort of device, although the energetic loss-rate adheres to different, more complex, laws. However, currently, heat-pipes only cover relative short distances up to ~30 kilometer (Kavvadias & Quoilin, 2018) and (excess) heat by industry or power plants is used only in the direct vicinity. In the future, this might change as the economic feasibility of long distance (over 100 kilometer) heat transmission increases, as pointed out in work of Kavvadias and Quoilin (2018). One of the potential reasons contributing to this result can be attributed to the fact that thermal storage is significantly cheaper than electrical storage. In this work however, the assumption is made that heat providing devices and heat demands are connected to the same thermal node, making the modelling of heat pipes redundant. Instead, for district heating, a fixed heat loss of 10% is assumed and directly applied to the 'Thermal load' device by increasing the demand accordingly. This loss-rate was earlier described in the work of Lämmle (2020).

3.1.4 Storage

In contrast to connecting devices that are leveling energy prices among nodes by acting as spatial arbitragers, storage devices can be seen as temporal arbitragers. These devices can store energy at times when prices are low and discharge when prices are high. Therefore, these devices dampen price volatility. The storage device has no explicit cost-function but is constrained by maximum (dis-)charging levels and maximum storage levels. This work considers perfect storage devices with no leakage effects. In future work an ambient temperature dependent penalty term could be added to model efficiency decreases in cold environments. The constraints are summarized in Equation (3.9). The variable name SoC stands for 'State of Charge' and represents the battery level. The parameters CR_{max} and DR_{max} represent the maximum hourly charging rate and dis-charging rate respectively.

$$\begin{aligned} p &\geq DR_{max} \\ p &\leq CR_{max} \end{aligned} \Bigg\} \quad \rightarrow \textit{Charging constraints}$$
$$\begin{aligned} SoC &\geq 0 \\ SoC &\leq SoC_{max} \end{aligned} \Bigg\} \quad \rightarrow \textit{Battery level constraints} \tag{3.9}$$
$$SoC^t - SoC^{t-1} = p \quad \rightarrow \textit{Change in battery level should equal the input}$$

3.1.5 Dissipation devices

Dissipation devices are devices with zero cost-function that can absorb or dissipate energy, whether it be electrical of thermal, in times of excess supply. Therefore, they ensure not only the feasibility of the system but also keep track on how much free electricity or useful heat is wasted. They are subjected to only one constraint, namely, they can not provide energy to the system:

$$\begin{aligned} p &\geq 0 \\ q &\geq 0 \end{aligned} \Bigg\} \quad \rightarrow \textit{Constraints for electrical and thermal energy dissipation respectively} \tag{3.10}$$

Note. Instead of defining the objective functions based on operational cost, these functions could alternatively be related to GHG emissions when the objective is to minimize the total emissions of the energy model rather than the operational cost.

3.2 Problem Formulation

The previous section explained how an energy system can be modelled as a network consisting of nodes, edges and devices. Figure 3.1 shows a schematic representation of such an energy network. With the additional information from Section 3.1, such a network can also be depicted as a bipartite graph, as illustrated in Figure 3.5.

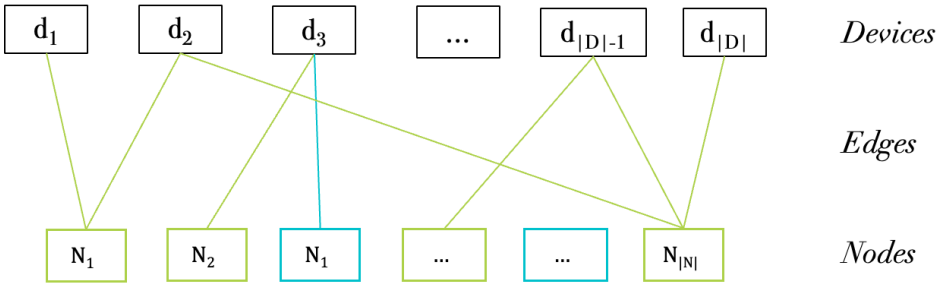

Figure 3.5: *Schematic representation of a device-based energy model using a bipartite graph in which device 2 represents a transmission line (it is connected to two nodes), and device 3 represents a CHP plant (it is connected to an electrical node and a thermal node)*

One of the main objectives in this thesis is concerned with investment strategies for future energy networks. However, identifying 'good' investments is not straightforward and depends on the perspective taken. For example, transmission lines can reduce price differences between two regions by transporting energy. From the perspective of the generator on the 'expensive' side, this could be disadvantageous, as it may reduce revenue. Similarly, France blocked the expansion of the European energy network between Spain and Germany, aiming to protect its own nuclear sector from the competition posed by cheaper renewable energy generation in Spain (Oliver, 2014). This work will focus on investments from the perspective of the public sector, meaning that investments should benefit the public, either by minimizing energy prices or reducing GHG emissions. In this thesis, the focus will be on minimizing energy prices. Summarizing, the objective is to find the optimal set of investments, such that it minimises the overall energy cost while satisfying demand. This problem is formulated in Problem (3.11), with the explanation of the variables and sets is given at the top of the next page.

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i \in D \cup I} f_i(x_i, p_i^d) \\
\text{subject to} \quad & p_{ij}^e = \left(p_i^d\right)_j && \forall (ij) \in E && \rightarrow \textit{Energy transport on edge (ij) equals} \\
& && && \textit{energy sent from (potential) device i to node j} \\
& p_j^n = \frac{1}{|E_j|} \sum_{(ij) \in E} p_{ij}^e = 0 && \forall j \in N && \rightarrow \textit{Energy balance at node j} \\
& A_i x_i + B_i p_i \leq C_i && \forall i \in D && \rightarrow \textit{Internal constraints of device i} \\
& A_i x_i + B_i p_i \leq z_i C_i && \forall i \in I && \rightarrow \textit{Internal constraints of potential device i} \\
& \sum_{i \in I} z_i \cdot c_i \leq B && && \rightarrow \textit{Budget constraint} \\
& z_i \in \{0, 1\} && \forall i \in I &&
\end{aligned}
\tag{3.11}
$$

Symbol	Description
D	Set of existing devices
E	Set of edges
$E_i \subset E$	set of edges connected to device i
$E_j \subset E$	set of edges connected to node j
I	Set of investments (also called potential devices or candidate devices)
N	Set of nodes
R	Set of regions
k	Iteration indicator
p_i^d	Energy flow corresponding to device i : $\in \mathbb{R}^{T \times N}$
p_{ij}^e	Energy flow on edge (ij) connecting device i with node j : $\in \mathbb{R}^T$
p_j^n	Energy imbalance at node j : $\in \mathbb{R}^T$
λ_{ij}^e	Dual variable on edge (ij) connecting device i with node j : $\in \mathbb{R}^T$
x_i	Internal variable(s) belonging to device i
A_i	Constraint matrix belonging to variable x_i
B_i	Constraint matrix belonging to variable p_i
$(M)_c$	Indicator of c^{th} column of matrix M

The objective function in Formulation (3.11) minimizes the total operational cost. Although it sums over the set of existing devices and the set of investments, this does not imply that all investments participate in the network. The investments are limited in their operation when they are not selected. This is ensured by the fourth constraint, where z_i indicates whether the investment is selected. The first constraint ensures that energy transported on an edge equals the amount of energy exchanged between the device and the node. The second constraint ensures an energy balance at each node. The internal constraints for all existing devices are captured by the third constraint. Furthermore, the fifth constraint ensures that the total investment cost does not exceed the budget. The final constraint makes sure that no partial investment is part of the solution.

Formulation (3.11) also clarifies why the problem can, conceptually, be viewed as two optimization problems. The minimum operational cost for a combination of investments corresponds to a unique flow schedule, which can be found by optimizing the OEF problem. Indirectly, two questions need to be answered: which devices should be selected and what is the corresponding optimal flow schedule? One strategy of solving this problem is to repeatedly fix a combination of investments and solve the corresponding OEF problem. A sophisticated version of this approach is the *Branch & Bound* algorithm. This algorithm solves combinatorial optimization problems by systematically evaluating all meaningful combinations. Applied on Problem (3.11), this means that branching is done on each investment, and the objective at each node is found by solving the corresponding OEF problem. Nodes are pruned when the lower bound of the node is higher than the existing upper bound. While this method is often used in commercial solvers for solving (Mixed) Integer Programming problems, the number of possible combinations grows

exponentially with the number of investment options, making this method potentially inefficient. Alternatively, Section 3.4 proposes two heuristics that evaluate only parts of the solution space but still rely on repeatedly solving the OEF problem. Therefore, it would beneficial to solve this problem efficiently, which will be the topic of Section 3.3.

Lower & Upper Bound

The upper and lower bound can be interpreted as the maximum and minimum operational costs of satisfying energy demand over a full-year period, given a certain budget. The initial upper bound can be easily defined as the benchmark model with only the existing devices. The lower bound is calculated by adding all potential devices to the benchmark model and relaxing the binary investment decision variables of Formulation (3.11) to continuous variables. Although the solution may include, for example, installation of fractional wind turbines, it gives a theoretical lower bound against which we can evaluate the performance of the heuristics.

3.3 Optimal Energy Flow

This section describes how the OEF problem can be solved using distributed optimization techniques for an energy model as described in Section 3.1. Ignoring the investments decisions from Problem (3.11) results in the optimization that is concerned with the optimal energy flow schedule for a fixed set of devices. This problem is summarized below:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i \in D} f_i(x_i, p_i^d) \\
\text{subject to} \quad & p_{ij}^e = \left(p_i^d\right)_j && \forall (ij) \in E && \rightarrow \textit{Energy transport on edge (ij) equals} \\
& && && \quad \textit{power sent from device i to node j} \\
& A_i x_i + B_i p_i \leq C_i && \forall i \in D && \rightarrow \textit{Internal constraints of device i} \\
& p_j^n = \frac{1}{|E_j|} \sum_{(ij) \in E} p_{ij}^e = 0 && \forall j \in N && \rightarrow \textit{Energy balance at node n}
\end{aligned}
\tag{3.12}
$$

Comparing Problem (3.12) and (2.10) shows their similarity, and indeed, the OEF problem can be seen as a 'Market Clearing Problem', which was earlier described by Kazempour (2024). The significance of this result is that the same 'Exchange ADMM' algorithm can be used to optimize Problem (3.12). The energy balance constraint at each node acts here as the global constraint, linking multiple devices. These constraints are therefore relaxed and included in the Augmented Lagrangian function. Having moved the global balancing constraint to the objective function enables the use of the Exchange ADMM algorithm, as summarized in Algorithm 2 and described in Section 2.1. The algorithm shows that the devices can be optimized in parallel due to their independence from each other, revealing one of the biggest advantages of distributed optimization. Furthermore, the second step in the algorithm, calculating the imbalance at the nodes, can also be parallelized because the nodal power balance is independent of other nodes. However, due to the low complexity of the calculation, the time savings will probably not be as

significant.

Algorithm 2 Exchange ADMM Algorithm with Device Decomposition

Initialize $p_{ij}^{e,0}, p_i^{d,0}, p_j^{n,0}, \lambda_{ij}^{e,0}, \epsilon$
$k \leftarrow 0$
$\mu \leftarrow \epsilon + 1$ ▷ Set μ to an arbitrary number above ϵ for the first iteration
while $\mu \geq \epsilon$ **do**
  for each $i \in D$ **do** ▷ For each Device i
    $p_i^{d,k+1} \leftarrow \text{argmin}_{p_i^d} L_\rho(p_i^{d,k}, p_{ij}^{e,k}, \lambda_{ij}^{e,k} \mid \forall j : (ij) \in E)$ ▷ Device optimization
    $p_{ij}^{e,k+1} = \left(p_i^{d,k+1}\right)_j$
  end for
  for each $j \in N$ **do** ▷ For each Node j
    $p_j^{n,k+1} \leftarrow \frac{1}{|E_j|} \sum_{(ij)\in E} p_{ij}^{e,k+1}$ ▷ Equilibrium calculation
    $\lambda_{ij}^{e,k+1} \leftarrow \lambda_{ij}^{e,k} + \rho \cdot p_j^{n,k+1}$ ▷ Dual update
  end for
  $\mu \leftarrow \|\lambda^{k+1} - \lambda^k\|$
  $k+ = 1$
end while

Dual variable interpretation

Section 2.1 explained that in Problem (3.12) there exists a Nash equilibrium as long as the problem is convex (proof given by Banks and Duggan (2004)). Within this equilibrium, no participant (such as a generator, battery, or other device) can deviate from this point without negatively impacting their own objective. The optimal strategy corresponding to this result is linked to duality theory, which states that the optimal energy price is equal to the dual variable of the balance constraint at each node. The link between dual variables and game theory is further discussed in work of Bradley, Hax and Magnanti (1977, Chapter 4.9). This result provides insights to the magnitude, spatial variability and temporal variability of the dual variables. For example, the energy price difference between two nodes is equal to the difference of their dual variables. When this difference is large, a new transmission line could level these prices to the point that the transmission cost (see Section 3.1.3) exceed the nodal price difference. Alternatively, the dual variable could behave volatile over time. This could imply that there is enough capacity of renewable energy, but that there is no cheap alternatives when renewable sources are unavailable. Batteries can decrease this price volatility by storing 'cheap' renewable energy, which increases the overall utilization of renewables and decreases the average energy price. In other words, even without advanced modelling and optimization, the dual variables offer insights into potential improvements that could benefit the system.

3.4 Investment Optimization

This section describes the optimization related to determining the optimal investment strategy. By ignoring the decision variables and constraints related to the optimal energy flow from Problem (3.11), the objective shifts solely to selecting the optimal set of investments. This problem is summarized below:

$$
\begin{aligned}
\text{minimize} \quad & g(z) && \rightarrow \textit{Minimizing cost} \\
\text{subject to} \quad & \sum_{i \in I} z_i \cdot c_i \leq B && \rightarrow \textit{Budget constraint} \\
& z_i \in \{0,1\}, \qquad \forall i \in I
\end{aligned}
\tag{3.13}
$$

In this formulation, z represents the vector of decision variables, where the i^{th} element indicates whether investment i is selected. The objective function $g(z)$ represents the operational cost corresponding to this decision vector, while the cost of investment i is indicated by c_i. With a bit of creativity, Formulation (3.13) can be seen as a knapsack problem. However, the difficulty in solving this particular problem lies in the fact that the utility is not a constant but rather an unknown function dependent on all the other decision variables. This makes Problem (3.13) at least as hard as the knapsack problem. The next subsections will introduce two heuristics that require fewer computational resources than a *Branch & Bound* algorithm, and hopefully still provide high-quality solutions.

3.4.1 Dynamic Programming Heuristic

The difficulty of Problem (3.13) lies in the fact that the utility of each investment depends on the set to which it belongs, rather than being constant. When simplifying the problem and ignoring this fact, Problem (3.13) reduces to a 0/1 Knapsack problem which can be efficiently optimized using a Dynamic Programming algorithm. This is exactly the approach used in this heuristic. Each investment is added individually to the existing set of devices, and the OEF problem is solved for each scenario. Then, the effect (i.e. reduction of operational cost) corresponding to each investment is stored. The biggest disadvantage of this approach is the implicit assumption that investment X has no influence on the profitability of investment Y and vice-versa, while it is known this is not the case. An outcome of this assumption could be that the found solution might invest excessively in one technology or in one region. To partly overcome this, restrictions are imposed such that at each node, only one investment can be made for each type of technology (e.g., generation by renewable sources and storage devices). For example, at node j it is allowed to install a battery and a wind turbine, but not two batteries of different size or, PV **and** wind turbine(s). For transmission lines this restriction is not imposed. To achieve this, investments are categorized in certain subgroups. For example, sub-group 1 contains the following four investment possibilities:

- Install 3 MW capacity wind turbine at node 1
- Install 6 MW capacity wind turbine at node 1
- Install 500 kW capacity PV at node 1
- Install 4 MW capacity PV at node 1

Then, from each subgroup, at most one investment can be selected. This is a special variant of the knapsack problem that is also known as the 'Multiple-Choice Knapsack Problem' (MCKP) (Kellerer, Pferschy & Pisinger, 2004). The dynamic programming table corresponding this variant is schematically depicted in Figure 3.6, and the pseudo-code is summarized in Algorithm 3.

Algorithm 3 Pseudo-code for Dynamic Programming Algorithm solving MCKP

Initialize investments ▷ List of tuples (cost, utility, subgroup)
    budget ▷ Total cost should not exceed this value
    dynamic programming table ▷ Size is $(|I|+1)$ x $(B+1)$
for each $i \in I$ **do** ▷ For each investment i
    Solve OEF of benchmark model including investment i
    Save the objective value related to investment i in Table T
end for
Sort Table T based on **subgroup**
for each i **in range 1 to** $|I|$ **do**
    for each b **in range 1 to** B **do**
        Determine the optimal selection from the investments numbered 1, 2,..., i
        when budget b is available, and store the value in the DP-table ▷ See Fig. 3.6
    end for
end for
Use **backtracking** to find the optimal solution

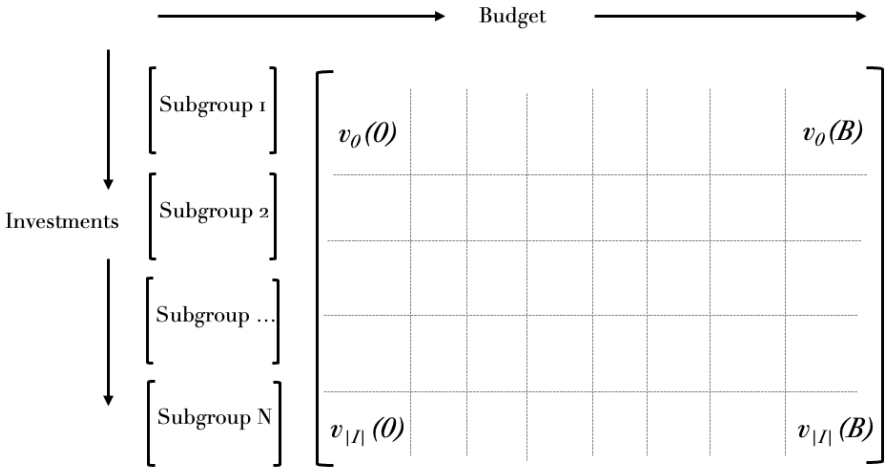

Figure 3.6: *Schematic figure of the Dynamic Programming Table corresponding to the MCKP problem.*

3.4.2 Relax & Fit Heuristic

This section describes the 'Relax & Fit' heuristic. The approach is as follows, the relaxed version of Problem (3.11) is solved and the (fractional) solution is retrieved. Each fractional investment gets assigned a fitness-score, the closer the investment is to an existing option, the higher the fitness-score. The investment with the highest fitness-score is selected and fixed within the model. The investment cost of this device is then subtracted from the (residual) budget, and the problem relaxation is solved again. This process continues until the remaining budget is insufficient to make any further investments. Assumed is that within this approach, the fitness-score represents the probability of a particular investment being part of the optimal solution. The heuristic is summarized in Algorithm 4.

Algorithm 4 Pseudo-algorithm for Relax & Fit Heuristic

Initialize set of investments I
    total budget B
    $s \leftarrow 0$ ▷ Amount of money spent
    $\mu \leftarrow 1$ ▷ Identifier if any investment could still be added in last iteration
    $S \leftarrow \emptyset$ ▷ Set of selected investments
while $\mu > 0$ **do**
    $\mu \leftarrow 0$
    Solve the relaxation of Problem 3.11
    for each $i \in I$ **do** ▷ For each investment i in the set of Potential Investments
        $i \leftarrow$ *fitness-score*
    end for
    Sort set I by *fitness-score*
    for each $i \in I$ **do**
        if $c_i \leq B - s$ **then** ▷ If investment cost is still within the residual budget
            Fix investment i
            Add investment i to S
            Remove investment i from I
            $s \leftarrow s + c_i$
            $\mu \leftarrow 1$
            break
        end if
    end for
end while

3.5 Network example

To evaluate the ADMM algorithm for solving the OEF problem and to test the proposed heuristics, a 12-bus energy model with both electrical and thermal nodes is constructed. Figure 3.7 illustrates the model, where thick black lines represent electrical nodes, thick blue lines indicate thermal nodes, and thin black lines between electrical nodes denote transmission lines. The

dashed lines are potential transmission lines that are part of the investment options. The overall distribution of energy demand between the residential sector, the manufacturing sector, and the commercial sector is approximately even, with each sector accounting for around one-third of the total electrical load. This approximation is in line with values reported by *U.S. Electricity System About the U.S. Electricity System and its Impact on the Environmen* (n.d.). Furthermore, approximately 30% of all households are connected to 'District Heating', and another 30% of households have a connection to a heat pump. For convenience, only the thermal demand of households is considered, and not those of commercial buildings. The network consists of nodes with different characteristics (e.g. predominantly households, manufacturing industry, commercial buildings, or power generation). Some of these nodes are connected and some of them could be connected in the future when the connection is part of the investment options. There are conventional as well as renewable electrical generators installed. The network from Figure 3.7, without the investment options, is referred to as the benchmark model. The data sources for the various load profiles and the availability of renewable energy are summarized in Table 3.2. Additionally, Tables 3.3, 3.4, and 3.5 provide details on the existing devices and transmission lines in the benchmark model. Figure 3.8 shows the cost functions of the three fossil-fuel-based generators plotted against their electrical output. Tables 3.6 and 3.7 list all the investment options for this model. Next, the investment costs were adjusted to ensure that all technologies contribute approximately equally to the solution of the linear relaxation of Problem (3.11). This adjustment aims to balance the cost-effectiveness of each technology, preventing any single technology from dominating the optimal solution. These costs do not necessarily reflect the capital expenditures found in reality.

Table 3.2: *Data sources for weather statistics and load profiles*

Device	**Data**	**Source**
Renewable generators	Availability of Wind or Solar	Renewables Ninja [Pfenninger and Staffell (2016), Staffell and Pfenninger (2016)]
Thermal Load	Thermal Load Profiles	Renewables Ninja
Electrical Load	Load Profiles	BDEW [Hellwig (2003)] ELMAS [Bellinguer et al. (2023)]

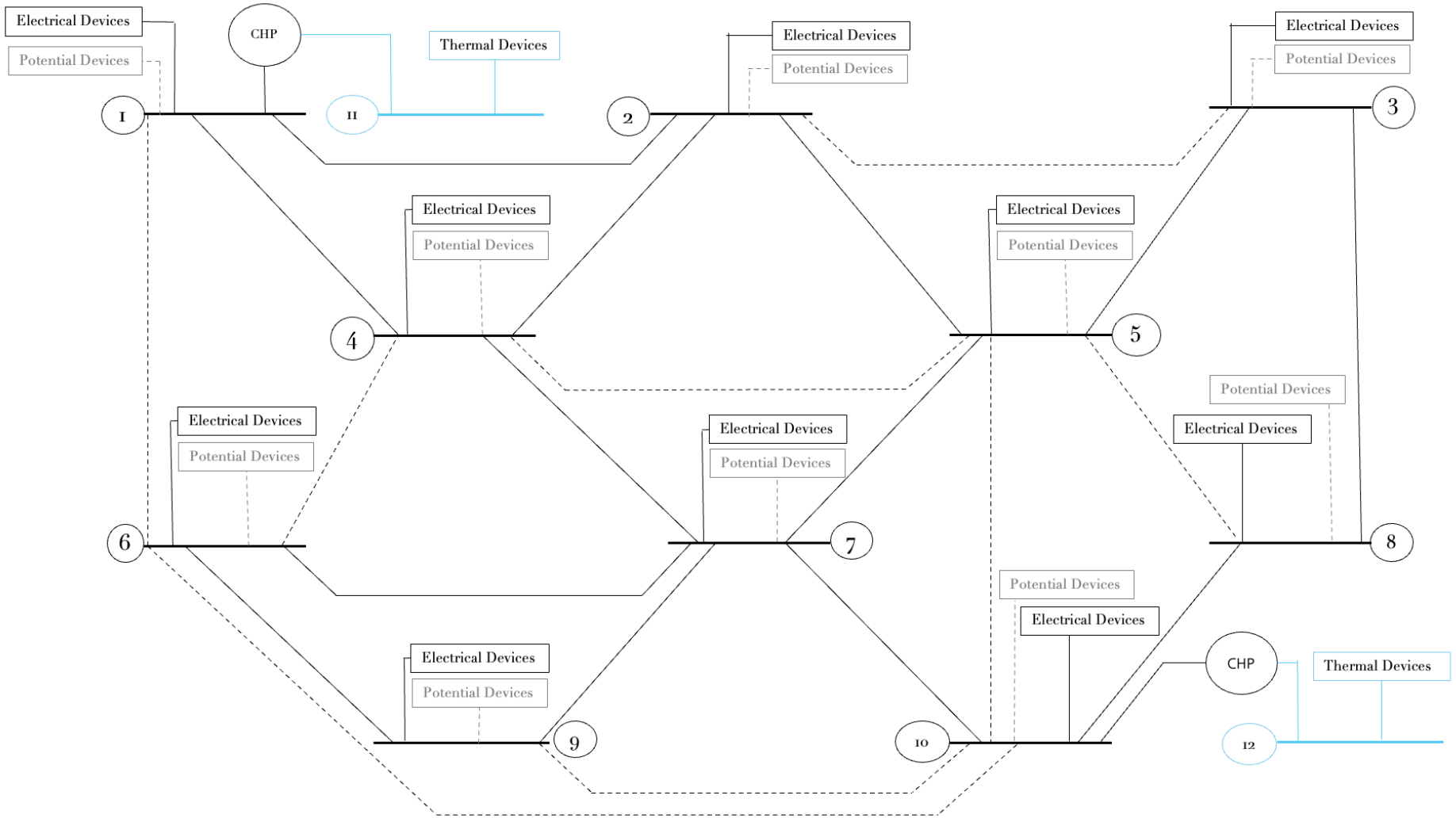

Figure 3.7: *The benchmark model consisting of* 10 *electrical nodes (indicted by thick black lines) and* 2 *thermal nodes (indicated by the thick blue lines). Thin black lines connecting electrical nodes represent the transmission lines. Thin dashed lines represent transmission lines that are not installed, but are part of the set of investment options. The numbers specify the node number. The specifications of the devices installed at each node, as well as the specifics of the investment options, can be found in the subsequent tables.*

Table 3.3: *Specifications of the benchmark model*

Net	Type	Device	Capacity [MW(h)] Load [$\frac{\text{GWh}}{\text{a}}$]	Ramping constr. [$\frac{\text{MW}}{\text{h}}$] (Dis)charging constr. [$\frac{\text{MW}}{\text{h}}$]	Cost $f(p)$
Net 1	Electrical	Power Dissipation	∞	-	-
		Electrical Load (Commercial)	35	-	-
		Electrical Load (Households)	30	-	-
		Electrical Load (Heat-pumps)	7.63	-	-
		CHP	18	1	$(p-3)^2 + 40 * p + 20 * q + 1$
Net 2	Electrical	Power Dissipation	∞	-	-
		Electrical Load (Households)	5	-	-
Net 3	Electrical	Power Dissipation	∞	-	-
		Electrical Load (Manufacturing)	20	-	-
		Battery	60	3	-
Net 4	Electrical	Power Dissipation	∞	-	-
		Electrical Load (Households)	6	-	-
		Battery	30	2	-
Net 5	Electrical	Power Dissipation	∞	-	-
		Generator	20	0.5	$0.75 * (p-4)^2 + 33 * p + 1$
		Renewable (PV)	11	-	-
Net 6	Electrical	Power Dissipation	∞	-	-
		Renewable (PV)	6	-	-
		Renewable (Wind)	12	-	-
Net 7	Electrical	Power Dissipation	∞	-	-
		Electrical Load (Manufacturing)	5	-	-
		Electrical Load (Commercial)	4	-	-

Table 3.4: *Continuation of Table 3.3*

Net	Type	Device	Capacity [MW(h)] Load $[\frac{\text{GWh}}{\text{a}}]$	Ramping constr. $[\frac{\text{MW}}{\text{h}}]$ (Dis)charging constr. $[\frac{\text{MW}}{\text{h}}]$	Cost $f(p)$
Net 8	Electrical	Power Dissipation	∞	-	-
		Electrical Load (Commercial)	4	-	-
		Electrical Load (Households)	10	-	-
		Renewable (PV)	2	-	-
Net 9	Electrical	Power Dissipation	∞	-	-
		Electrical Load (Manufacturing)	45	-	-
		Renewable (Wind)	6	-	-
Net 10	Electrical	Power Dissipation	∞	-	-
		CHP	12	2	$0.25 * (p-2)^2 + 43 * p + 22 * q + 1$
		Electrical Load (Commercial)	25	-	-
		Electrical Load (Households)	20	-	-
		Electrical Load (Heat-pumps)	5.09	-	-
Net 11	Thermal	Heat Dissipation	∞	-	-
		Thermal Load	25.68	-	-
Net 12	Thermal	Heat Dissipation	∞	-	-
		Thermal Load	17.12	-	-

Table 3.5: *Specifications of the transmission lines of the benchmark model*

Transmission Line nr.	From Net	To Net	Capacity [MW]	Cost $f(p)$
1	1	4	2	$0.25 \cdot p^2$
2	1	2	2	$0.25 \cdot p^2$
3	2	5	2	$0.25 \cdot p^2$
4	3	5	2	$0.25 \cdot p^2$
5	3	8	2	$0.25 \cdot p^2$
6	4	7	3	$0.25 \cdot p^2$
7	5	7	4	$0.25 \cdot p^2$
8	6	9	4	$0.25 \cdot p^2$
9	6	7	4	$0.25 \cdot p^2$
10	7	9	4	$0.25 \cdot p^2$
11	7	10	4	$0.25 \cdot p^2$
12	8	10	4	$0.25 \cdot p^2$

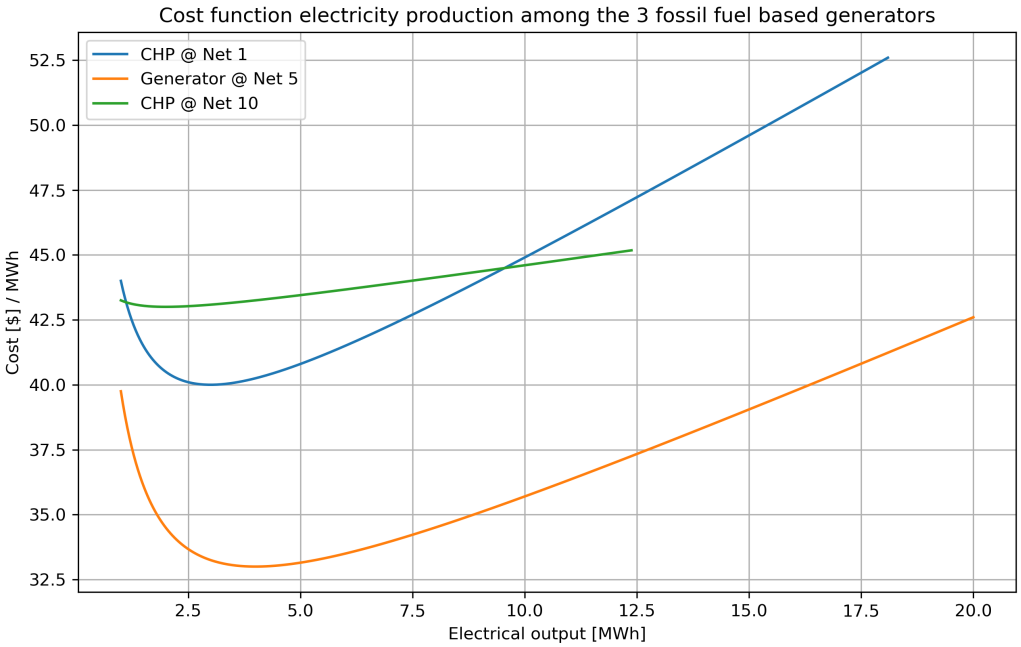

Figure 3.8: *The cost function of the three fossil-fuel based generators in the benchmark model*

Table 3.6: *Investment options (excluding transmission lines) in the benchmark model*

Device	Capacity [MW(h)]	Possible at Nodes	Cost [10^4\$]
Renewable (Wind)	3	2 - 10	55
	6	2 - 9	100
Renewable (PV)	1	1 - 10	20
	4	2 - 9	38
Battery	30	2 - 10	15
	60	2 - 10	28

Table 3.7: *New transmission lines as part of the investment options for the benchmark model*

Transmission Line	From Node	To Node	Length	Capacity [MW]	Cost [10^4\$]
Line 1	1	6	*Long*	2	15
				4	25
Line 2	2	4	*Long*	2	15
				4	25
Line 3	2	3	*Short*	2	10
				4	15
Line 4	4	6	*Long*	2	15
				4	25
Line 5	4	5	*Short*	2	10
				4	15
Line 6	5	10	*Long*	2	15
				4	25
Line 7	5	8	*Long*	2	15
				4	25
Line 8	6	10	*Long*	2	15
				4	25
Line 9	9	10	*Short*	2	10
				4	15

4 | Results

This chapter will present the numerical results for solving Problem (3.11) based on the example given in Section 3.5. First, Section 4.1 presents the main results for the optimal operation of the benchmark model (i.e. the example model with no investments). Subsequently, Section 4.2 compares the performance of distributed optimization versus centralized optimization (i.e. non-decomposed optimization model optimized by a commercial solver) for the OEF problem. Lastly, Section 4.3 will discuss the results of the two heuristics as described in Section 3.4 for solving the investment optimization problem.

4.1 Benchmark Example

Solving the OEF problem corresponding to the benchmark model results in an optimal energy flow schedule. The key results of this schedule, namely, operational cost, the amount of curtailed electricity from renewable resources, and the amount of unused usable thermal energy are summarized in Table 4.1. The yearly average dual variable at each node and the standard deviation of the dual variable, based on 8760 hourly observations (representing one full year), are presented on the left side of Table 4.2. The mean dual variable indicates the average nodal energy price, while the standard deviation is listed as a proxy for price volatility. The right hand side of Table 4.2 summarizes information regarding the operation of the transmission lines. The left column lists the amount of time that each line operates at its maximum capacity. This can be an indicator how often, and where potential congestion occurs in the grid, which on its turn can sustain price differences between nodes. The right column shows the average utilization of the transmission lines, defined as the amount of energy transferred in the full year as a percentage of the maximum energy that can be transferred on that line. The largest number in each column of Table 4.2 is indicated in a bold font, and blue entries indicate that it concerns a thermal node. Noticeable is the large standard deviation of the dual variable at node 6. This can be explained by the lack of electricity demand and the sole presence of renewable generation at this node. Consequently, the marginal cost of energy at node 6 is highly dependent on the availability of solar radiation and wind, leading to considerable price fluctuations and a high standard deviation of the dual variable. Furthermore, the transmission lines connecting node 6 to node 7 (line 9), node 6 to node 9 (line 8), and node 7 to node 10 (line 11), often operate at their maximum capacity. The transmission lines connected to node 6 are crucial for transporting 'cheaply' generated renewable energy from this node, and the fact that these lines frequently operate at their limit could indicate grid congestion issues in this area.
For illustrative purposes, Figure 4.1 shows the energy flow at node 6 over ten consecutive days.

The diurnal pattern of solar power generation is clearly observable. It is also noticeable that a significant amount of power is dissipated when both solar and wind power are generated. This can occur due to a lack of demand, insufficient storage capacity, and/or insufficient capacity of the transmission network.

Table 4.1: *Main findings of the benchmark model after solving corresponding the OEF problem.*

Objective value [$]	Curtailed Electricity [MWh]	Unused Usable Heat [MWh]
4,736,034	19,194.26	46,140.54

Table 4.2: *Overview of yearly average nodal pricing of electricity and thermal energy in the left table. The right table shows the percentage of time transmission lines operate at their maximum capacity and the amount of energy transferred in the full year as a percentage of the maximum. The maximum values in each column are indicated with a bold font. Blue entries indicate that it concerns a thermal node.*

Node nr.	Mean λ	Std. Dev. λ
1	41.81	18.41
2	29.30	18.70
3	**44.61**	6.01
4	37.90	12.21
5	27.97	18.10
6	23.98	**21.68**
7	33.08	16.74
8	42.49	8.74
9	32.77	17.73
10	41.92	8.93
11	3.11	6.94
12	2.45	6.41

Line nr.	Operating at max. capacity [%]	Average utilization [%]
1	2.08	63.65
2	0.00	86.16
3	6.79	50.16
4	0.00	**94.96**
5	1.50	50.86
6	0.00	50.61
7	16.19	56.17
8	27.34	77.25
9	38.06	69.57
10	0.43	40.69
11	**39.13**	68.57
12	0.00	32.20

4.2 Distributed optimization vs. Centralized optimization

One of the research questions is whether an OEF model can benefit from distributed optimization. The *'Exchange ADMM'* algorithm is applied to evaluate the computational time for solving three different OEF problems varying in size and time horizons. These results are presented in Table 4.3. Here, the 1-bus system corresponds to the benchmark model including only the first node. In the 3-bus system, nodes one, two, and four are included along with their relevant transmission lines. Lastly, in the 12-bus system the full benchmark model is considered. Additionally, either 1 or 8760 time-steps are considered, corresponding with the first hour of 2019 and the full year respectively.

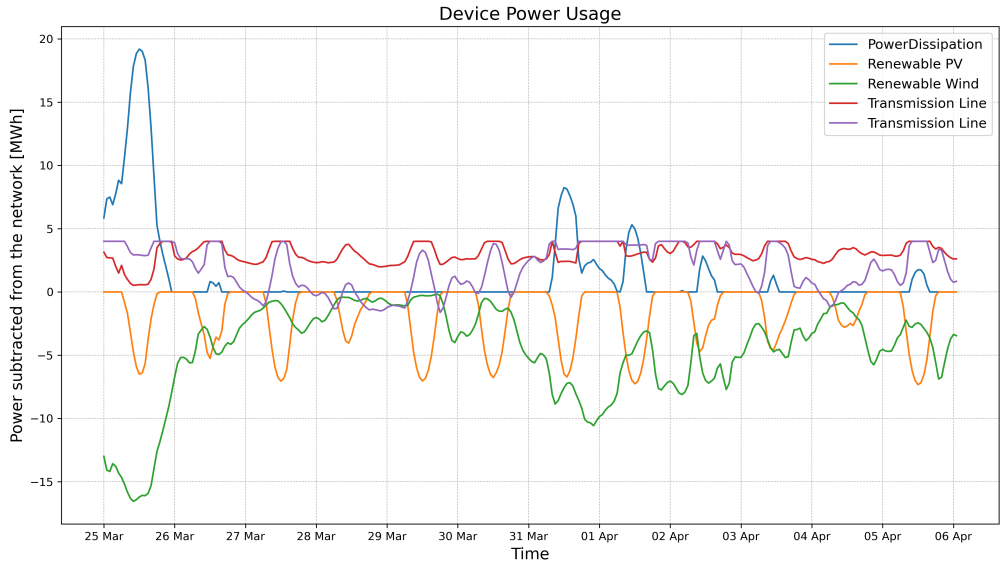

Figure 4.1: *For illustrative purposes, results of the incoming and outgoing power on node* 6 *of the benchmark model, between March,* 25^{th} *and April,* 6^{th}*. Positive/negative values correspond to subtracting/providing power from/to the node.*

Table 4.3: *Results of the comparison of a centralized optimization approach and the distributed ADMM method. The values in last column are based on the example model from Section 3.5.*

	1-bus		3-bus		12-bus	
Time-steps	1	8760	1	8760	1	8760
Centralized						
Computational Time$[s]$	$3.82 \cdot 10^{-4}$	$9.07 \cdot 10^{-2}$	$5.15 \cdot 10^{-4}$	$3.18 \cdot 10^{-1}$	$2.43 \cdot 10^{-3}$	2.08
ADMM						
Computational Time$[s]$	$3.23 \cdot 10^{-2}$	460.43	$1.15 \cdot 10^{-1}$	1122.43	14.36	> 3600

The ADMM results should be interpreted as follows: they represent the computational time when the problem is not parallelized, despite parallelization being one of its major advantages. Assuming the algorithm dedicates all its computational time to device optimization, with negligible time spent on the other operations, and assuming all devices require an equal amount of optimization time, and an unlimited number of computer cores are available, the computational time could potentially be reduced by a factor equal to the number of devices.
Even if we account for the potential parallelization, for these small to medium sized problems, the ADMM approach is not as efficient as a centralized approach solved using a commercial solver. ADMM would be most beneficial for cases where the full model is too large to solve, while each subproblem is much smaller. This is not the case in the presented example model, but this could be the case in real-life models. When applying ADMM, it is important to also optimize the penalty parameter ρ due to its significant effect on the convergence rate.

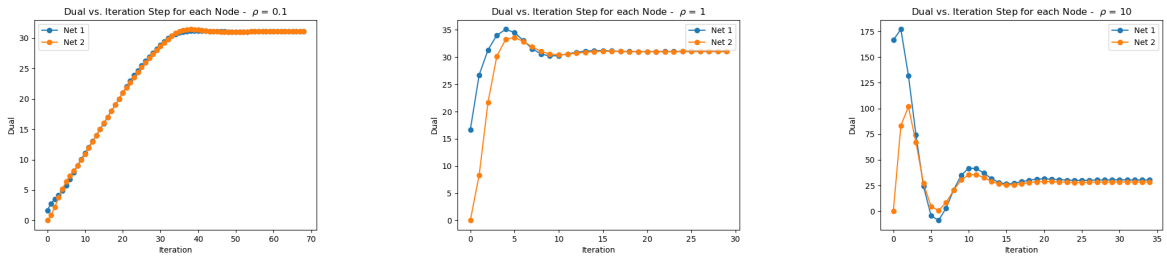

Figure 4.2: *Magnitude of the dual variables versus the iteration number for different sizes of ρ. The left figure corresponds to $\rho = 0.1$, while the middle and right figure use 1 and 10 respectively. Higher penalty parameters are characterized with smaller lead-times, but with higher overshoot.*

Penalty Parameter

Figure 4.2 plots the dual variable versus the iteration number for nodes 1 and 2 of the example model when only one time-step is considered. This relation is plotted for three different values of ρ, and the relation between the size of the penalty parameter and the lead-time and overshoot can be clearly seen. Selecting the optimal penalty parameter is a well-known problem in control theory, where the objective is to minimize lead-time and overshoot. To achieve this, it is often beneficial to take larger steps at the beginning of the optimization and reduce the step size as the optimal solution is approached. Instead of a fixed penalty term, a dynamically decreasing term through the iterations could be introduced. Alternatively, the penalty term could be set proportionally to the primal residual of the linking constraint at each node (i.e., the power imbalance). In that case, the penalty term varies not only across iterations but also differs across nodes based on their respective power imbalances.

4.3 Heuristics

The second research question is concerned with efficient methods for identifying the optimal investment combination. As previously mentioned, the number of combinations increases exponentially with the number of investment options, and exact methods require too much computational resources. Problem (3.11) did not terminate in reasonable time using a commercial solver, emphasizing the dependence on heuristic methods. Therefore, we do not discuss the numerical performance of the commercial solver further. In this section the results of the *Relax & Fit* heuristic and the *Dynamic Programming* heuristic are presented based on their performance with respect to the benchmark model and a budget of 240 [10^4 \$]. The main findings are presented in Table 4.4 below. Sections 4.3.1 and 4.3.2 provide a detailed analysis of the results obtained from the respective heuristics.

Table 4.4: *A comparison of the different methods used to solve the investment problem related to the benchmark model from Section 3.5.*

Relaxed Problem	
Upper bound [$]	4,736,034
Lower bound [$]	3,077,615
Curtailed Electricity [MWh]	9,180.88
Unused usable Heat [MWh]	10,872.87
Dynamic Programming	
Computational Time [s]	568.08
Objective Value [$]	3,695,707
Gap to LB [-]	16.72%
Curtailed Electricity [MWh]	5,069.23
Unused usable Heat [MWh]	19,196.37
Relax & Fit	
Computational Time [s]	212.32
Objective Value [$]	3,181,253
Gap to LB [-]	3.26%
Curtailed Electricity [MWh]	9,399.23
Unused usable Heat [MWh]	12,508.25

4.3.1 Dynamic Programming

Based on the benchmark model, the investment set corresponding to the solution of the *Dynamic Programming* heuristic is summarized in Table 4.5. This table lists all the devices that are added next to the already existing ones in the benchmark model. This selection resulted in a gap of 16.72% with respect to the lower bound found by solving the relaxation of Problem (3.11). It significantly reduced the curtailment of electricity from renewable resources by nearly 75% compared to the benchmark model. Despite outperforming the *Relax & Fit* heuristic in reducing curtailment, it resulted in more wasted thermal energy. Additionally, the objective function is considerably higher than for the *Relax & Fit* heuristic. The selected investments seem to over-expand in grid transmission capacity. This could be explained by the fact that many transmission lines reduce the objective value considerably, but the algorithm does not account for the fact that, as connectivity grows, the marginal gain of each additional transmission line decreases. The relative large investment in grid capacity and electrical storage reduces spatial price differences, temporal price differences, and electricity curtailment. However, these technologies only facilitate the utilization of 'cheap' renewable energy but do not increase the overall generation. Looking at the left-hand side of Table 4.6, it can be observed that the mean value of the dual variable and the standard deviation of the dual variable at the different nodes are, apart from node 1, indeed very similar. Furthermore, the right-hand side of Table 4.6 shows that the transmission lines hardly operate at their maximum capacity anymore, indicating little to none congestion issues.

Table 4.5: *Selected investments based on results of the Dynamic Programming heuristic*

Device	Capacity [MW(h)]	(From) Node	To Node
Transmission Line 13	2	9	10
Transmission Line 14	4	6	10
Transmission Line 15	2	5	8
Transmission Line 16	4	5	10
Transmission Line 17	2	4	5
Transmission Line 18	2	4	6
Transmission Line 19	2	2	3
Storage	30	2	-
Storage	30	5	-
Storage	30	6	-
Storage	30	7	-
Storage	30	9	-
Wind Turbine	3	3	-

Table 4.6: *Information regarding the nodal prices and the operation of the transmission lines corresponding to solution of the Dynamic Programming heuristic, on the left and right side respectively. Bold transmission lines indicate the newly installed connections that are part of the solution of the Dynamic Programming heuristic. Blue entries indicate that it concerns a thermal node.*

Node nr.	Mean λ	Std. Dev. λ
1	39.83	26.47
2	28.06	13.06
3	28.28	12.97
4	28.24	12.99
5	27.82	12.77
6	27.21	13.63
7	28.22	13.09
8	28.32	12.92
9	28.57	13.35
10	28.51	13.05
11	3.77	7.30
12	8.68	8.43

Line nr.	Operating at max. capacity [%]	Average utilization [%]
1	0.11	74.32
2	0.00	80.83
3	0.00	36.63
4	0.00	44.88
5	0.00	22.95
6	0.00	7.94
7	0.06	24.15
8	1.32	55.28
9	1.30	37.99
10	0.00	20.15
11	0.00	15.68
12	0.00	15.44
13	0.57	33.43
14	5.46	51.65
15	1.28	50.46
16	0.17	35.78
17	0.00	41.75
18	0.00	63.58
19	0.00	22.64

4.3.2 Relax & Fit Results

Based on the benchmark model, the *Relax & Fit* heuristic outperformed the *Dynamic Programming* heuristic by selecting the set of investments summarized in Table 4.7. This resulted in a gap of 3.26% with respect to the lower bound of Problem (3.11). Additionally, it resulted in less curtailment of electricity (a reduction of approximately 50%) and less wasted thermal energy (a reduction of more than 70%) compared to the result of the benchmark model (see Table 4.1). Based on these results, it appears that the heuristic is effective in identifying which technologies should be installed at which nodes to reduce the overall operational costs of the system. The dual variables of the benchmark model, as presented in Table 4.2, provide additional insight to this solution. A high-capacity transmission line is installed between two nodes where the difference of the yearly average dual variables was initially among the highest in the model. Furthermore, storage technologies are installed at nodes where the standard deviations of the dual variable were originally the highest and fourth highest, indicating increased price volatility at those nodes. Lastly, generators based on renewable energy sources are installed at three nodes that rank in the top four in terms of the average dual variable in the benchmark model, indicating high marginal costs at these nodes. These results are in line with the interpretation of the dual variables, given in Section 3.3.

Looking at the left-hand side of Table 4.8, the system corresponding to the results of the *Relax & Fit* heuristic indicates an average reduction in the marginal cost of energy at all electrical nodes and a decrease in the standard deviation of the marginal cost at most electrical nodes. On the contrary, both thermal nodes (nodes 11 & 12) experience price increases and probably increased price volatility compared to the model without investments. This can be explained by the reduced dependency on CHPs to meet electrical demand, resulting in less 'free' excess thermal energy being available. The right-hand side of Table 4.8 shows that, compared to the benchmark model, the investment strategy significantly reduces the frequency at which transmission lines 8, 9, and 11 operate at their maximum limit. This reduction indicates fewer congestion issues and, consequently, less spatial variability of energy prices. Additionally, the average utilization of each transmission line, except for line 12, is lower. This suggests that the added transmission capacity makes the system more robust, as most lines have, on average, more reserve capacity.

Table 4.7: *Selected investments based on results of the Relaxed & Fit heuristic*

Device	Capacity [MW(h)]	(From) Node	To Node
Transmission Line 13	2	4	5
Transmission Line 14	4	1	6
Wind Turbine	3	3	-
Wind Turbine	6	8	-
Storage	30	5	-
Storage	30	6	-
PV	2	1	-

Table 4.8: *Information regarding the nodal prices and the operation of the transmission lines corresponding to solution of the Relax & Fit heuristic, on the left and right side respectively. Bold transmission lines indicate the newly installed connections that are part of the solution of the Relax & Fit heuristic. Blue entries indicate that it concerns a thermal node.*

Node nr.	Mean λ	Std. Dev. λ
1	30.20	15.03
2	26.76	15.59
3	28.66	15.24
4	28.11	15.10
5	26.37	14.89
6	26.64	16.29
7	28.14	15.41
8	28.83	15.58
9	28.35	15.99
10	30.08	14.78
11	7.89	9.17
12	8.20	9.27

Line nr.	Operating at max. capacity [%]	Average utilization [%]
1	0.02	51.27
2	0.00	80.32
3	6.12	54.30
4	0.00	71.50
5	8.58	48.90
6	0.00	24.50
7	13.28	48.94
8	4.27	63.17
9	6.50	49.36
10	0.08	30.76
11	12.60	59.50
12	4.19	40.38
13	0.00	57.51
14	0.00	70.42

5 | Conclusion & Discussion

This work considered a budget constrained energy system expansion problem. This optimization problem is modeled as an MIQP problem, and the possibility of solving it using distributed optimization in combination with heuristics is explored to keep the problem computationally tractable. To assess the research questions, a medium-sized energy model is designed, consisting of an electrical network and a thermal network interconnected by CHP plants. Solving the OEF problem for this model using the distributed optimization algorithm *'exchange ADMM'* did not provide computational advantages compared to solving the same problem centrally with the commercial solver *'Gurobi'*. Although the optimization contained over $100,000$ decision variables, the commercial solver was extremely efficient in finding the optimal objective value. When the problem size would significantly increase, distributed optimization could become noticeably faster compared to commercial solvers or even a necessity.

Two heuristics were assessed for identifying the (sub-)optimal set of investment options: the *Dynamic Programming* heuristic and the *Relax & Fit* heuristic. Based on the presented model and the data used, the *Relax & Fit* heuristic performed best, while the commercial solver could not solve the model in a reasonable time. The heuristic achieved an operational cost that was only 3.26% higher than the lower bound, which was determined by solving the relaxation of the same problem and allowed for the installation of fractional investment options. The optimal set of investment options has a strong connection with the dual variables (i.e. the marginal cost of energy) defined at each node. In the solution found by the *Relax & Fit* heuristic, transmission lines are installed between nodes whose initial dual variables differed relatively much. Increased connectivity levels energy prices between nodes, and overall, enables the utilization of the cheapest energy form available in the network. Furthermore, storage facilities are installed at nodes where, initially, the standard deviation of the dual variable was high, a metric that was used as a proxy for price volatility. These investments store 'cheap' renewable energy and help absorb price shocks when energy from renewable sources is insufficient. This reduces price volatility and increases the utilization of clean energy. Lastly, generators that run on renewable resource are installed at nodes where generation was otherwise relatively expensive, which aligns with the fact that power originating from renewable sources is generated with low operational cost. Therefore, the characteristics of the dual variables, derived from the nodal pricing system, can provide additional insights into an energy system and could help indicating technologies that lower the overall operational cost.

5.1 Recommendations for *'District'*

This thesis is written in collaboration with the ISE department at the Fraunhofer-Gesellschaft research institute where they developed *'District'*, a regional sector-coupling energy model designed to provide reliable insights into how a future energy network could look when constrained by emission regulations and/or plans for expanding heat and electrical demand (Thomsen, 2017). Because large projects require extensive computational resources, the team is interested in whether decomposition techniques and distributed optimization algorithms can benefit their operation. Although the current structure of *'District'* does not easily allow decomposition by device, it is suitable for decomposition by area or region. These regions consist of generating devices, storage facilities, and electrical loads and are connected to other regions by transmission lines and heat pipes. It is recommended to *cut* these connections to create smaller sub-problems, and place a node to connect the two loose ends, as shown in Figure 5.1.

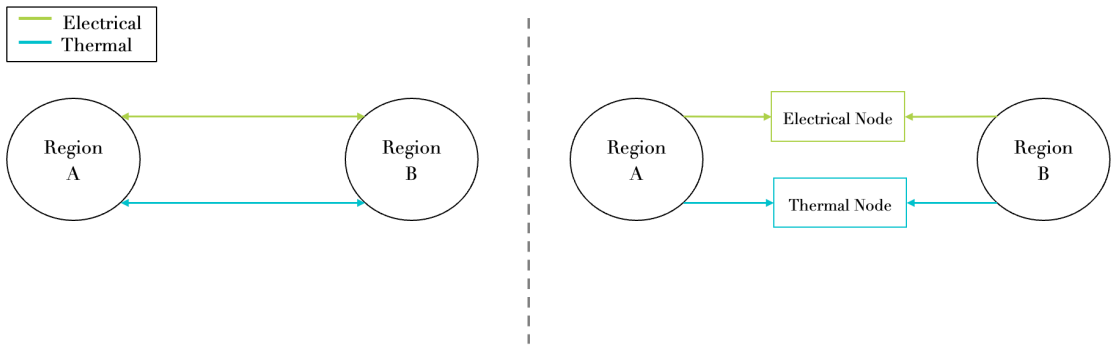

Figure 5.1: *Schematic representation on how the District model could be decomposed.*

From here the same algorithm as in Section 3.3 can be used. In fact, the two decomposition techniques result in the same problem, with just different interpretations. In regional decomposition, the coupling constraint always involves only two energy flows (i.e. energy subtracted/ supplied to the two regions at each end of the transmission line or heat pipe). Consequently, transmission lines are not considered devices but function as nodes. The regions can be viewed as large devices that are composed of multiple devices grouped together.

However, Chapter 4 shows that distributed optimization might not be the holy grail for solving the *'District'* model faster for large projects. The binary or integer decision variables pose a real challenge regarding whether, and in what time, the problem will terminate, especially when multiple technologies have a competitive value and branching algorithms can not easily prune. This problem will not disappear when using a distributed optimization approach. Moreover, it does not guarantee convergence to the global optimum. For larger projects, it is recommended to relax the integrality constraints and use a heuristic (e.g., Relax & Fit) to construct the (sub)-optimal solution. Parallelized distributed optimization could become valuable when projects become so large that the commercial solver struggles to find a solution even for the relaxed problem.

5.2 Economic limitations

Monopolies

In Chapter 2 it is argued that generators cannot unilaterally deviate from a certain strategy while simultaneously improving their profit, which corresponds to energy prices being equal to the marginal cost. This result is based on the implicit assumption of a competitive market, however, since we are dealing with a physical product, this may not hold when the transmission network is inadequate. Consequently, certain nodes could be relatively isolated, allowing generators to exploit their market power.

Inelastic Energy Demand

In the example network presented in Section 3.5, fixed load profiles are used to model the energy use of different sectors. These load profiles are based on empirical research, where end-users very likely did not have dynamic energy contracts. However, the model presented in Chapter 3 applies a pricing system that does adhere to dynamic pricing. Therefore, load profiles might not be as inelastic as presented when there is an incentive to use energy at times prices are low. The results presented in Chapter 4, which demonstrate potential savings for energy users from certain investments, rely on a benchmark model and benchmark prices. These benchmark prices are influenced by inelastic demand profiles. However, if energy purchasers were to observe dynamic nodal prices rather than contract prices, their behaviour could change, potentially altering the benchmark itself.

5.3 Technical limitations

Whether Problem (3.11) would be solved by the exact *Branch & Bound* algorithm or a heuristic, it is unavoidable that the OEF problem or the linear relaxation of Problem (3.11) needs to be evaluated repeatedly. It turned out that out-of-the-box solvers are very efficient in solving this problem on the presented model. Based on reasoning alone, it could be argued that the implementation of parallel computing with the ADMM algorithm was not going to boost computational efficiency to a level where it would outperform the commercial solver. Therefore, the effect of parallelization is not further explored, and its impact remains uncertain. In relation to this, it is still unclear at what problem size distributed optimization becomes faster than commercial solvers.

5.4 Future research

As described in Section 3.1, CHP plants interconnect electrical and thermal networks. It is shown that these devices do not significantly increase the complexity of the model and allow for optimization of the integrated system. Future research could expand on this idea by incorporating a hydrogen network. It would be interesting to see how a hydrogen plant can be modelled as a convex device, how the operation of the integrated system changes, and if the dynamic nodal pricing system could be used to predict profitability. Additionally, it would be interesting

to see how energy load profiles would change when energy users see dynamic nodal prices instead of fixed contractual prices. With customers incentivized to use electricity during periods of lower prices, research could be done how this changes behavioural patterns with respect to energy usage. Lastly, the distributed optimization algorithm performed worse than expected, probably due to the size of the OEF problem. It would be interesting to determine the problem size at which commercial solvers begin to struggle with optimizing convex problems, and when distributed optimization becomes a faster alternative.

References

Banks, J. & Duggan, J. (2004). *Existence of nash equilibria on convex sets* (Tech. Rep. No. WP20). University of Rochester - Wallis Institute of Political Economy. Retrieved from https://ideas.repec.org/p/roc/wallis/wp20.html

Bastianel, G., Ergun, H. & Hertem, D. V. (2022, OCT). Linearised optimal power flow problem solution using dantzig - wolfe decomposition. *2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, *Not available*, Not available. Retrieved from https://dx.doi.org/10.1109/isgt-europe54678.2022.9960529 doi: 10.1109/isgt-europe54678.2022.9960529

Bellinguer, K., Girard, R., Bocquet, A. & Chevalier, A. (2023, OCT). Elmas: a one-year dataset of hourly electrical load profiles from 424 french industrial and tertiary sectors. *Scientific Data*, *10*, Not available. Retrieved from https://dx.doi.org/10.1038/s41597-023-02542-z doi: 10.1038/s41597-023-02542-z

Benders, J. F. (1962, DEC). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, *4*, 238–252. Retrieved from https://dx.doi.org/10.1007/bf01386316 doi: 10.1007/bf01386316

Boyd, S. (2010). Distributed optimization and statistical learning via the alternating direction method of multipliers. Retrieved from https://dx.doi.org/10.1561/9781601984616 doi: 10.1561/9781601984616

Bradley, S. P., Hax, A. C. & Magnanti, T. L. (1977). *Applied mathematical programming*. Addison Wesley Publishing Company.

Chatzivasileiadis, S. (2018). *Lecture notes on optimal power flow (opf).* Retrieved from fdsafdsafs

Durvasulu, V. & Hansen, T. M. (2018, NOV). Market-based generator cost functions for power system test cases. *IET Cyber-Physical Systems: Theory amp; Applications*, *3*, 194–205. Retrieved from https://dx.doi.org/10.1049/iet-cps.2018.5046 doi: 10.1049/iet-cps.2018.5046

Dvorkin, V., Kazempour, J., Baringo, L. & Pinson, P. (2018, DEC). A consensus-admm approach for strategic generation investment in electricity markets. *2018 IEEE Conference on Decision and Control (CDC)*. Retrieved from https://dx.doi.org/10.1109/cdc.2018.8619240 doi: 10.1109/cdc.2018.8619240

Hellwig, M. (2003). *Entwicklung und anwendung parametrisierter standard-lastprofile* (Unpublished doctoral dissertation). Technischen Universität München.

Holt, C. A. & Roth, A. E. (2004, MAR). The nash equilibrium: A perspective. *Proceedings of the National Academy of Sciences*, *101*, 3999–4002. Retrieved from https://dx.doi.org/

10.1073/pnas.0308738101 doi: 10.1073/pnas.0308738101

Joyce, P. (1984, undefined). The walrasian tatonnement mechanism and information. *The RAND Journal of Economics*, *15*, 416. Retrieved from https://dx.doi.org/10.2307/2555449 doi: 10.2307/2555449

Kavvadias, K. C. & Quoilin, S. (2018, APR). Exploiting waste heat potential by long distance heat transmission: Design considerations and techno-economic assessment. *Applied Energy*, *216*, 452–465. Retrieved from https://dx.doi.org/10.1016/j.apenergy.2018.02.080 doi: 10.1016/j.apenergy.2018.02.080

Kazempour, J. (2024). *Renewables in electricity markets.* Retrieved from -

Kellerer, H., Pferschy, U. & Pisinger, D. (2004). The multiple-choice knapsack problem. *Knapsack Problems*, 317–347. Retrieved from https://dx.doi.org/10.1007/978-3-540-24777-7_11 doi: 10.1007/978-3-540-24777-7_11

Kraning, M. (2014). Dynamic network energy management via proximal message passing. Retrieved from https://dx.doi.org/10.1561/9781601987150 doi: 10.1561/9781601987150

Krause, T., Andersson, G., Fröhlich, K. & Vaccaro, A. (2011, JAN). Multiple-energy carriers: Modeling of production, delivery, and consumption. *Proceedings of the IEEE*, *99*, 15–27. Retrieved from https://dx.doi.org/10.1109/jproc.2010.2083610 doi: 10.1109/jproc.2010.2083610

Lämmle, M. (2020). Measurement results of a district heating system with decentralized incorporated solar thermal energy for an energy, cost effective and electricity grid favourable intermitting operation. *Proceedings of the ISES EuroSun 2020 Conference – 13th International Conference on Solar Energy for Buildings and Industry*. Retrieved from https://dx.doi.org/10.18086/eurosun.2020.02.06 doi: 10.18086/eurosun.2020.02.06

Moehle, N., Busseti, E., Boyd, S. & Wytock, M. (2019). *Large scale optimization in supply chains and smart manufacturing.* Springer International Publishing. Retrieved from https://dx.doi.org/10.1007/978-3-030-22788-3 doi: 10.1007/978-3-030-22788-3

Moehle, N., Shen, X., Luo, Z.-Q. & Boyd, S. (2019, MAY). A distributed method for optimal capacity reservation. *Journal of Optimization Theory and Applications*, *182*, 1130–1149. Retrieved from https://dx.doi.org/10.1007/s10957-019-01528-5 doi: 10.1007/s10957-019-01528-5

Oliver, C. (2014). *Franco-spanish energy spat tests eu.* Retrieved 2024-05-28, from https://www.ft.com/content/8e94079c-585f-11e4-b331-00144feab7de

Pfenninger, S. & Staffell, I. (2016, NOV). Long-term patterns of european pv output using 30 years of validated hourly reanalysis and satellite data. *Energy*, *114*, 1251–1265. Retrieved from https://dx.doi.org/10.1016/j.energy.2016.08.060 doi: 10.1016/j.energy.2016.08.060

Staffell, I. & Pfenninger, S. (2016, NOV). Using bias-corrected reanalysis to simulate current and future wind power output. *Energy*, *114*, 1224–1239. Retrieved from https://dx.doi.org/10.1016/j.energy.2016.08.068 doi: 10.1016/j.energy.2016.08.068

Thomsen, J. (2017, NOV). Enhancing operation of decentralized energy systems by a regional economic optimization model district. *Energy Systems*, *9*, 669–707. Retrieved from

https://dx.doi.org/10.1007/s12667-017-0261-9 doi: 10.1007/s12667-017-0261-9

U.S. Electricity System about the u.s. electricity system and its impact on the environmen. (n.d.). https://www.epa.gov/energy/about-us-electricity-system-and-its-impact-environment. (Accessed: 2024-05-16)

Xu, D., Wu, Q., Zhou, B., Li, C., Bai, L. & Huang, S. (2020, OCT). Distributed multi-energy operation of coupled electricity, heating, and natural gas networks. *IEEE Transactions on Sustainable Energy*, *11*, 2457–2469. Retrieved from https://dx.doi.org/10.1109/tste.2019.2961432 doi: 10.1109/tste.2019.2961432

Zhang, X., Shahidehpour, M., Alabdulwahab, A. & Abusorrah, A. (2015, SEP). Optimal expansion planning of energy hub with multiple energy infrastructures. *IEEE Transactions on Smart Grid*, *6*, 2302–2311. Retrieved from https://dx.doi.org/10.1109/tsg.2015.2390640 doi: 10.1109/tsg.2015.2390640