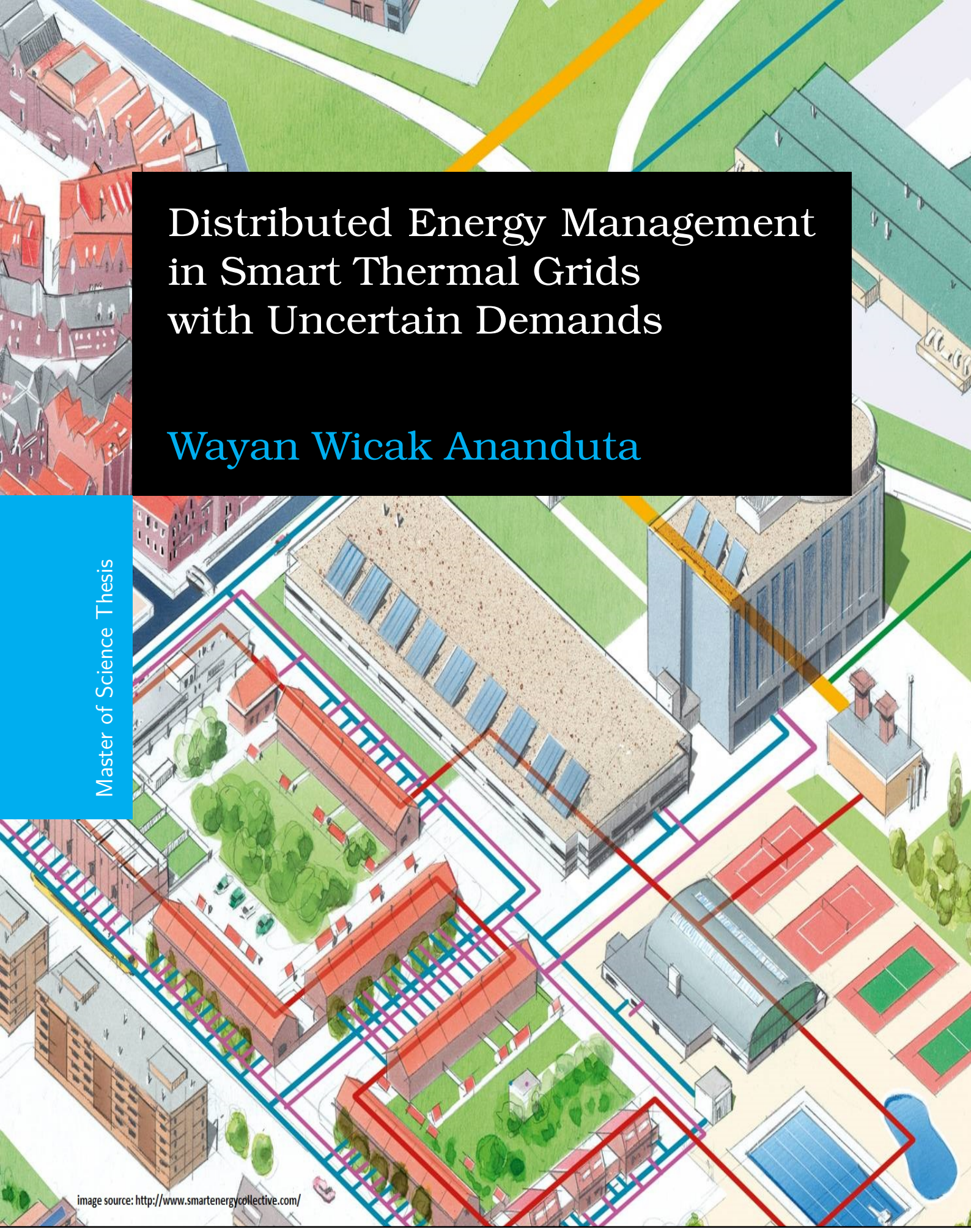
Distributed Energy Management in Smart Thermal Grids with Uncertain Demands

Wayan Wicak Ananduta

Master of Science Thesis



TU Delft Delft University of Technology

Delft Center for Systems and Control

Distributed Energy Management in Smart Thermal Grids with Uncertain Demands

Master of Science Thesis

For the degree of Master of Science in Systems and Control at Delft University of Technology

Wayan Wicak Ananduta

30 August 2016

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology

Delft University of Technology
Department of
Delft Center for Systems and Control (dcsc)

The undersigned hereby certify that they have read and recommend to the Faculty of Mechanical, Maritime and Materials Engineering (3mE) for acceptance a thesis entitled

Distributed Energy Management in Smart Thermal Grids with Uncertain Demands

by

Wayan Wicak Ananduta

in partial fulfillment of the requirements for the degree of

Master of Science Systems and Control

Dated: 30 August 2016

Supervisor(s):

Vahab Rostampour, M.Sc

Dr. ir. Tamás Keviczky

Reader(s):

Dr. ir. Matthijs Spaan

Prof. dr. ir. Hans Hellendoorn

Abstract

Smart Thermal Grids (STGs) represent a new concept in the energy sector that involves the use of the smart grid concept in thermal energy networks connecting users, such as households, buildings and greenhouses, to each other via a transport line of thermal energy. In this concept, there exists an energy management system that aims to improve the efficiency, reliability and sustainability of the energy production and the distribution of energy. This highlights the necessity of constructing a high level control unit which sets the operating points of the production units such as boilers, micro Combined Heat Power (CHP) generators, and chillers for every agent.

In this thesis, we develop a framework for the energy management which incorporates the model of the smart thermal grid and the energy demand profile of the agents. This framework is based on a Model Predictive Control (MPC) strategy in which the grid is addressed as a large-scale uncertain system. We formulate a mixed-integer chance-constrained optimization problem for the planning of the operation of the production units for all agents in the grid in the presence of uncertain thermal energy demand profile. In order to deal with the chance constraints together with integer variables, the robust randomized method in [1], which was particularly developed for this problem, is employed. This technique allows us to handle mixed-integer problems and stochastic programming in a unified framework and provide a-priori probabilistic guarantees for the obtained solution. Motivated by the need for a more flexible and scalable framework, we then extend this method to distributed computation schemes using the Alternating Direction Method of Multipliers (ADMM). The resulting performance enhancement in terms of the total operational costs observed in the simulation was substantial and comparable with the centralized control approach.

Finally, we investigate the opportunity of improving the efficiency of energy usage when seasonal storage systems exist in the thermal grids. We first propose a dynamic model for the seasonal storage systems and then, incorporate it in the energy management problem formulation. Due to the annual cyclic dynamical behavior of the seasonal storage systems, this leads to a multi-rate albeit very complex optimization problem. To this end, we develop a hierarchical MPC to solve such problem together with a discussion on the resulting optimization problems in a receding horizon setting. The technical developments were validated on a realistic benchmark problem (three-agent thermal grid in Utrecht, The Netherlands). The

simulation results show that the proposed method was able to provide better usage of the seasonal storage.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Firstly, I would like to thank my supervisors, Dr. ir. Tamás Keviczky and Vahab Rostampour, M.Sc. Without their guidance, immense knowledge and feedbacks, I would not be able to finish this thesis. More than that, I have learned many lessons about research which I really appreciate. I sincerely wish that we can keep in touch.

Secondly, I would like to thank my parents and Desy for always believing in and supporting me. Thank you for providing me with continuous encouragement throughout my years of study and through the process of writing this thesis. This accomplishment would not have been possible without them.

Finally, I would like to acknowledge Lembaga Dana Pengelolaan Pendidikan (LPDP) for providing me the scholarship for my master study in Delft University of Technology.

Delft, University of Technology
30 August 2016

Wayan Wicak Ananduta

Chapter 1

Introduction

1-1 Background

According to [2], up to 50% of the total global energy consumption is allocated to heat production, most of which is related to heating purposes in buildings or houses. This underlines not only the importance of thermal grids but also the necessity to improve them. The efficiency of the systems is one of the key aspects of such improvement and control engineering is seen as a promising solution. In fact, the notion of smart grids is introduced in conjunction with those that have the ability to autonomously manage their operations so that they are not only efficient but also reliable. In this thesis, we want to solve such control problem for thermal grids.

The main goal of these systems is obviously to fulfill the thermal energy demand. Thus information of the demand and possibly on future predictions is very important to the controller. Since this varies not only over time but also between consumers, we consider it as an uncertain variable so we would likely have a more realistic description of the systems. Robust and stochastic control approaches are available when uncertainties are present in the system. In this thesis we are more interested in exploring the latter because it could provide a less conservative solution, which would mean a better performance [3].

One of the major challenges that is faced in connection with this problem is the scale of these systems. They are viewed as large-scale systems since they consist of a large number of consumers and producers which are connected through distribution pipelines [4]. Moreover, nowadays besides centralized production units which serve many consumers in large geographical areas, there also exist decentralized ones which are owned by and serve only one of or a few consumers [4]. When dealing with some issues that could be found in the controlling of large-scale multi-agent systems, such as high computational demand, scalability, flexibility and safety, it is distributed control approaches that are preferred to the centralized one [5].

Despite increasing the complexity of the systems, the existence of decentralized production units is beneficial to the improvement of efficiency. Not only because they have higher efficiency than the centralized ones [6], [7] but also because they open up possibilities for the

consumers, which are the agents of the system, to play a bigger role in maintaining the energy balance of the whole system by exchanging thermal energy [1]. This needs to be taken into account when formulating the optimization problem.

In addition, other features and limitations of the components of thermal grids have to be considered and properly modeled. For instance, the existence of seasonal thermal energy storages could be exploited to improve the efficiency. Moreover, some of the operational constraints of each of the components should also be met. Due to the nature of the components or when the feature of exchanging thermal energy is present in the systems, there could be coupling between the agents which complicate the problem.

The problem that we wish to solve is determining the optimal operating points, in terms of efficiency and reliability, of the production units. This task requires designing a high level controller to manage the operation of the production units. In this regard, it is assumed that each production unit has a low level controller which is able to track the set points. A class of control methods which is suitable for the energy management system is Model Predictive Control (MPC). MPC is a model based controller design procedure in which the computation is based on solving an optimization problem. MPC computes control actions repeatedly by minimizing a cost function over a finite prediction horizon based on predictions obtained by a system model subject to the constraints that the problem has [8].

As previously mentioned, the control method has to be able to deal with the fact that thermal grids are uncertain coupled large-scale systems. On the one hand, a stochastic approach called the randomized method is considered to tackle the uncertainty. We are motivated to follow this approach due to the fact that it is a data-driven method and nowadays measurement data is widely available. Moreover, a-priori probabilistic guarantees of the solution can also be provided by this method [9]. On the other hand, in order to deal with the large-scale issue and the coupling between the agents, we investigate and develop suitable distributed control techniques that are based on dual decomposition. Furthermore, we also develop a hierarchical MPC scheme in order to improve the usage of seasonal storages. In the following section, some results of the previous research in MPC, in particular stochastic and distributed approaches are discussed.

1-2 Related Work

Considerable research in MPC methods has been carried out since the last decades. It includes the development of a stochastic approach called randomized MPC and distributed methods. A brief discussion on these methods and some of the literature results related to the energy balance problem is presented in this section.

Two randomized MPC methods were already proposed by Blackmore, *et al.* in 2003 and Batina in 2004, even though they both have huge drawbacks such as their high computational demand [10]. However, the work of [3], [11] and [9] introduces the scenario approach which is capable of tackling the drawbacks. The randomized MPC methods that are based on this were then presented by Schildbach, *et al* in [10], Matusko and Borrelli in [12], and Prandini in *et al* [13].

Some literature results which present the applications of the randomization-based approach to solving stochastic problems are [14], [15] and [16]. In [14], the randomized MPC is proposed to

regulate the building temperature.The main challenge to the implementation of this method is that the number of decision variables is too large so the authors proposed a distributed algorithm to reduce the computational burden which is based on the dual decomposition method. In [15], this method is implemented in an uncertain nonlinear system. Since the system is non-convex, the theoretical result of the lower bound does not apply to this system. Therefore the number of scenarios are perceived as tuning variables and the approach is viewed as a heuristic approach. In [16] the randomization-based approach is implemented to solve reserve scheduling problems for uncertain power systems. Even though the problem is formulated as a convex problem, the number of decision variables was too large which led to tractability issues. Thus, the randomization-based method is again considered as a heuristic approach.

In [17], Calafiore, *et al.* study the application of the scenario-based approach to the mixed-integer problem. They state that it is possible but computationally harder to solve than the continuous problem since the bound grows exponentially with respect to the decision variables. Further, in [18], Margellos, *et al.* state that the existence of binary variables in their problem infers that they cannot apply the procedure of producing a number of scenarios according to the scenario approach. In addition, the problem in [18] is a large-scale problem. Therefore in [18], the authors use the method proposed in [19] where a robust problem with bounded uncertainty is solved, where the bounds are computed using the scenario approach.

One of our main goals in this thesis is dealing with the issue arisen from the fact that thermal grids are large-scale systems. In order to cope with this, various distributed MPC methods, including the ones that are based on dual decomposition have been developed, as can be seen in some survey papers such as [8] and [20], or in [21]. The dual decomposition is an iterative method of solving the dual problem in which the Lagrange multipliers/prices which are assigned to the agents are adjusted during the iterations. A standard subgradient method is usually used to update the prices such as in [22], [23], while in [24], the convergence rate is improved by proposing an accelerated gradient method. Moreover, the work of [25] and [26] proposes updating techniques that are based on Newton's method. We provide an extensive explanation and discussion about this method in Appendix C.

However, the standard dual decomposition method can only deal with strictly convex problems. An extension of this which could solve a more general convex problems is called the Alternating Direction Method of Multipliers (ADMM) [27]. It uses the augmented Lagrangian to decompose the coupled system and it also has an iterative procedure of updating the primal and dual variables. Some articles which discuss ADMM-based distributed MPC methods are [28] and [29]. In [28] the authors propose this method in order to solve a multi-agent system which is coupled to the cost function. In addition, the authors of [29] provide a stopping criterion which could reduce the number of iterations at the expense of the optimality of the solutions.

The work of [1] provides a formulation of the energy balance problem in thermal grids which is a mixed-integer and chance-constrained problem and develops a stochastic MPC method, which is a mixture of robust and randomized techniques to solve it. Our main goal in this thesis is to extend the formulation and the method to a distributed setting. The authors of [30] and [31] has also addressed the energy balance problem of a grid. However, in both articles, they focus on electrical power even though in the latter article the thermal energy is also taken into account. Moreover, they do not consider the demand as an uncertain variable.

In these studies, they provide a distributed MPC method which is based on the standard dual decomposition method.

Hierarchical MPC methods that address multi-rate dynamic behavior can be found in [32], [33] and [34]. Both works of [32] and [33] are related to the energy management in buildings while the authors of [34] shows their hierarchical MPC approach applies to a chemical process system. The similarities between these works are as follows. The upper layer of the hierarchical MPC contains the slow rate dynamic to compute the steady state trajectories or set points and the lower layer in which the dynamic has faster rate is the controller that ensures that the system tracks the trajectory. Moreover, they also consider economic MPC problem in the upper layer.

1-3 Thesis Contribution

The contributions of this thesis are twofold:

1. We develop a distributed scheme as the extension of the unified framework for production planning of thermal grids which was developed in [1]. We decompose the problem, which is mixed-integer chance-constrained, based on the ADMM and provide two algorithms, one is fully distributed algorithm and the other one is with a central coordinator.
2. We explore the possibility of increasing efficiency by having seasonal storages in the grid. In this regard, we formulate the optimization problem that incorporates the dynamics and constraints of the seasonal storages. Furthermore, we develop a combination of hierarchical and distributed MPC scheme so that the annual cyclic behavior of the storages can be captured and better operating plans can be obtained while having distributed computations.

1-4 Report Structure

After this introductory chapter that explains motivations, related work, and the contributions of the thesis, we provide the distributed coordination of thermal grids with uncertain demands In Chapter 2. In this chapter, first we provide the mathematical modeling of thermal grids including the profile of the thermal energy demand of an agent. Afterwards, the day-ahead production planning problem is formulated. In these parts we extensively describe the cost function and constraints of the problem, as well as the compact reformulation that we used. We then explain the proposed algorithm to solve it which is called the robust randomized method and extend it to two distributed algorithms. In the last section of this chapter, we discuss the simulation study of the proposed methods.

Chapter 3 discusses the second case of the thermal grids where seasonal energy storages are present in the system. Here we discuss how they affect the systems and we formulate the corresponding production planning problems. Afterwards, we discuss the combination of hierarchical and distributed framework in which the robust randomized method is applied to tackle this problem. Lastly, we present and discuss the simulation study of this case.

We conclude this thesis in Chapter 4 by providing some remarks and recommendations for the future research in this topic.

Chapter 2

Distributed Coordination of Uncertain Thermal Grids

In this chapter we present a unified framework which solves a production planning problem of thermal grids. It is built based on the work of [1] which is extended to distributed settings. The framework presented in [1] is a mixed-integer chance-constrained problem and we propose two distributed robust randomized Model Predictive Control (MPC) methods which are based on the Alternating Direction Method of Multipliers (ADMM) to solve it.

This chapter begins with a discussion of the thermal energy demand profile and its stochastic model in Section 2-1. Afterwards, we explain the description of the system which includes the dynamics, constraints, and cost function in Section 2-2. It is followed by the formulation of the optimization problem in Section 2-3. Then, we present a tractable method to solve the problem in Section 2-4. Our main contribution in this part is the formulation of the distributed versions of the method which is described in Section 2-5. Finally, the simulation study in which we show the performance of the proposed algorithm is provided in Section 2-6.

2-1 Daily Thermal Energy Demand Profile

The profile of thermal energy demand is obtained using Low Energy Architecture (LEA) model in which it is viewed that the demand is only related to maintaining the indoor air temperature [35]. This model is principally based on the energy balance of the building and takes into account heat transfers that will affect the balance. Thus, the demand is the energy that is required to be supplied to or extracted from an agent (a building or house) so that the indoor air temperature is kept at the desired level. According to this model, the demand depends on environmental conditions, such as outside temperature, solar radiation and wind speed, as well as the insulation and the occupancy of the building. For instance, due to the difference between the desired indoor and outside air temperature, there is some energy that is exchanged between the building and the environment and the exchanges are considered as convective transfers which are related to the insulation of the building. When the outside

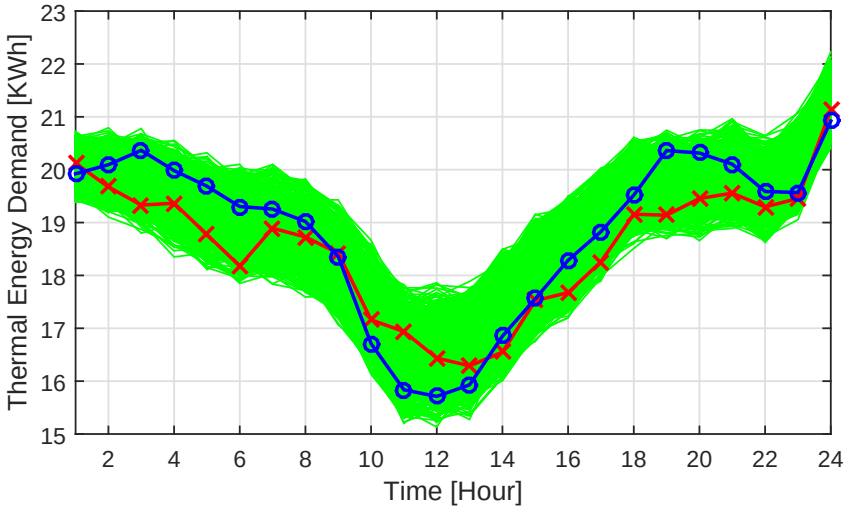


Figure 2-1: The real thermal energy demand (red line), the forecast demand (blue line) and 2000 scenarios that are generated from the Markov Chain model (green lines).

air temperature is lower than the desired indoor air temperature and since it is possible to transfer the energy through the wall, floor and roof, there will be some loss of energy which need to be compensated. When the situation is the other way around, the building gains thermal energy from the environment. Moreover, solar radiation, the use of lighting and equipments and the existence of people that occupy the building contribute to the thermal balance since they produce energy. In addition, we consider the demand is positive when the agent requires energy while it is negative when the agent has an excess of energy. A complete mathematical model is given in Appendix A-1.

In order to generate the forecast or random scenarios of the demand, we developed a stochastic version of the LEA model which was based on the Markov-Chain method [36], [37]. It was assumed that the uncertainty came from the outside air temperature and the occupancy of the building. Furthermore, the rest of the weather variables, *i.e.*, the wind speed and solar radiation, were assumed to be deterministic. Moreover, the transition matrix of the model was predefined. Figure 2-1 shows the plot of the real thermal energy demand of a building (red) as well as the forecast (blue) and 2000 random scenarios (green) that were generated from a Markov-Chain based model with 10 states. The value of the parameters related to the LEA model is given in Appendix A-2.

2-2 System Description of the Thermal Grid

In the thermal grids, we consider that each agent (a household or greenhouse) has its own decentralized production unit which consists of a boiler and a micro-Combined Heat Power (CHP)[1]. It is also viewed that an agent has the capability of exchanging thermal energy with each other. Moreover, it is also assumed that cooling demands are satisfied by using electrical

[1]Micro-CHPs are defined as the combined production of electrical (or mechanical) and useful thermal energy from the same primary source [7]

equipments and the electrical energy demands are fulfilled locally or by importing from a third party.

2-2-1 Dynamics of the System

Following [1], the problem that is dealt with is day-ahead production planning for all agents $i \in \mathcal{N} = \{1, 2, \cdots, N\}$, where N is the number of agents. For this problem, we consider a finite prediction horizon $N_h = 24$ with hourly steps and introduce k as the sampling time. First we define the main vector of the control variables of each agent, $\boldsymbol{u}_{m,i}(k) \in \mathbb{R}^{10}$, as

$$\boldsymbol{u}_{m,i}(k) = \left[p_{\mathbf{g},i}(k)\ p_{\mathbf{ug},i}(k)\ p_{\mathbf{dg},i}(k)\ h_{\mathbf{g},i}(k)\ h_{\mathbf{b},i}(k)\ h_{\mathbf{im},i}(k)\ u^{\mathbf{su}}_{\mathbf{g},i}(k)\ u^{\mathbf{su}}_{\mathbf{b},i}(k)\ z_{\mathbf{g},i}(k)\ z_{\mathbf{b},i}(k)\right]^T, \tag{2-1}$$

where $p_{\mathbf{g},i}(k)$ and $h_{\mathbf{g},i}(k)$ denote the electrical power and the thermal energy production of the micro-CHP, $p_{\mathbf{ug},i}(k)$ and $p_{\mathbf{dg},i}(k)$ denote the up and down spinning of electrical power by the micro-CHP, $h_{\mathbf{b},i}(k)$ denotes the amount of thermal energy that is produced by the boiler, $h_{\mathbf{im},i}(k)$ relates to the imported thermal energy from external parties, and $u^{\mathbf{su}}_{\mathbf{g},i}(k)$ and $u^{\mathbf{su}}_{\mathbf{b},i}(k)$ refer to the start-up cost variable of the micro-CHP and boiler. Moreover, $z_{\mathbf{g},i}(k)$ and $z_{\mathbf{b},i}(k)$ are the auxiliary variables that are required to model the minimum on and off time of each micro-CHP and boiler.

The grids are modeled based on the level of energy of each agent. It is coherent with the aim of the controller which is to ensure the efficiency and reliability of the energy production. In this regard, each agent is seen to have a thermal energy buffer. The dynamics of the energy level in the buffer, which indicates that the level at the next time instance depends on the current level, the energy produced, and the current demand, is represented in the following discrete-time equation:

$$h_{\mathbf{buf},i}(k+1) = \eta_{\mathbf{s},i}\left(h_{\mathbf{buf},i}(k) + h_{\mathbf{g},i}(k) + h_{\mathbf{b},i}(k) + h_{\mathbf{im},i}(k) + \sum_{j \in \mathcal{N}_{-i}} (1-\alpha_{ij}) h_{\mathbf{xc},ij}(k) - h_{\mathbf{d},i}(k)\right), \tag{2-2}$$

where $\eta_{\mathbf{s},i} \in (0,1)$ is the efficiency of the thermal energy storage and $h_{\mathbf{d},i}(k)$ is the thermal energy demand of agent i. Moreover, $h_{\mathbf{xc},ij}(k) \in \mathbb{R}$ denotes the thermal energy that is exchanged between agents i and j and α_{ij} denotes the transportation energy loss coefficient between those agents. Additionally, $\mathcal{N}_{-i}$ is defined as the set of neighbors of agent i as follows

$$\mathcal{N}_{-i} \subseteq \mathcal{N} \backslash \{i\}, \tag{2-3}$$

and all the exchanged energy of agent i with the neighbors, $h_{\mathbf{xc},ij} \forall j \in \mathcal{N}_{-i}$, are collected as an auxiliary vector $\boldsymbol{u}_{c,i}(k) \in \mathbb{R}^{|\mathcal{N}_{-i}|}$.

Furthermore, we introduce the imbalance error as the state of each agent. It refers to the difference between the energy level in the buffer ($h_{\mathbf{buf},i}(k)$) and the thermal energy demand ($h_{\mathbf{d},i}(k)$), as follows:

$$x_i(k) = h_{\mathbf{buf},i}(k) - h_{\mathbf{d},i}(k). \tag{2-4}$$

Having imbalance error as the state (x) of the system serves the purpose of having a variable which represents the reliability of the grid, which is one of the performance indicators in this study. We can easily say that the grid is reliable when the imbalance error is non-negative since it implies the supply is equal or bigger than the demand.

The dynamics of the state (imbalance error) can then be obtained by using its definition in Equation (2-4) and substituting $h_{\text{buf},i}(k+1)$ according to Equation (2-2). Thus we are able to form a linear dynamics equation of the state as follows:

$$x_i(k+1) = A_i x_i(k) + B_i \boldsymbol{u}_i(k) + H w_i(k), \tag{2-5}$$

where the input vector is $\boldsymbol{u}_i(k) = [\boldsymbol{u}_{m,i}^T(k)\ \boldsymbol{u}_{c,i}^T(k)]^T$ and $h_{\mathbf{d},i}(k+1)$ which is uncertain is considered to be the disturbance $w_i(k)$ of the system. Moreover, $A{=}\ \eta_{\mathbf{s}}$, $B{=}\ \eta_{\mathbf{s},i}[\boldsymbol{b}_1\ \boldsymbol{b}_{i,2}]$ where $\boldsymbol{b}_1 = [0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0]$ and $\boldsymbol{b}_{i,2} \in \mathbb{R}^{|\mathcal{N}_{-i}|}$ containing $(1-\alpha_{ij})\ \forall j \in \mathcal{N}_{-i}$ and $H{=}-1$. Note that the uncertain variable $w_i(k)$ is defined on a probability space Δ_i. It is assumed that Δ_i is endowed with the Borel σ-algebra and $\mathbb{P}$ is a probability measure defined over Δ_i.

Additionally, the status of the boilers and micro-CHPs is defined as a binary vector $\boldsymbol{v}_i(k) = \begin{bmatrix} v_{\mathbf{g},i}(k) & v_{\mathbf{b},i}(k) \end{bmatrix}^T \in \{0,1\} \times \{0,1\}$ which is defined as follows

$$\begin{aligned} v_{\mathbf{g},i}(k) &= \begin{cases} 0, & \text{micro-CHP is off,} \\ 1, & \text{micro-CHP is on,} \end{cases} \\ v_{\mathbf{b},i}(k) &= \begin{cases} 0, & \text{boiler is off,} \\ 1, & \text{boiler is on.} \end{cases} \end{aligned} \tag{2-6}$$

These binary variables are required to describe some of the constraints which are explained in the next subsection.

2-2-2 System Constraints

The constraints of the optimization problem is defined according to the operational condition or behavior of the system and the desired performance that has to be achieved. The following are the constraints that are taken into account.

Start-up Cost Constraints

The production unit consumes fuel to start up but during this time it does not produce energy. This behavior is modeled as startup cost constraints as follows

$$\boldsymbol{u}_i^{\mathbf{su}}(t) \geq \Lambda_i^{\mathbf{su}}(\boldsymbol{v}_i(t) - \boldsymbol{v}_i(t-1))\ , \quad c_{i,k}^{\text{su}} \geq 0, \quad t \in \{k, \cdots, k+N_h-1\}, \tag{2-7}$$

where Λ_i^{su} is a diagonal matrix in which the diagonal elements are the start-up costs of the micro-CHP and the boiler owned by an agent and t is the prediction time. Equation (2-7) indicates that every time the status of a boiler or micro-CHP changes from 0 to 1, which means it starts up, the additional cost of starting up is added to the agent.

Capacity Constraints

The capacity constraints of each agent $i \in \mathcal{N}$, for $t \in \{k, \cdots, k+N_h-1\}$ are defined in the following way:

$$v_{\mathbf{g},i}(t)p_{\mathbf{g},i}^{\min} \leq p_{\mathbf{g},i}(t) \leq v_{\mathbf{g},i}(t)p_{\mathbf{g},i}^{\max}\ , \tag{2-8}$$

$$v_{\mathbf{g},i}(t)h_{\mathbf{g},i}^{\min} \leq h_{\mathbf{g},i}(t) \leq v_{\mathbf{g},i}(t)h_{\mathbf{g},i}^{\max}\ , \tag{2-9}$$

$$v_{\mathbf{b},i}(t)h_{\mathbf{b},i}^{\min} \leq h_{\mathbf{b},i}(t) \leq v_{\mathbf{b},i}(t)h_{\mathbf{b},i}^{\max}\ , \tag{2-10}$$

$$h_{\mathbf{im},i}^{\min} \leq h_{\mathbf{im},i}(t) \leq h_{\mathbf{im},i}^{\max}\ , \tag{2-11}$$

$$h_{\mathbf{xc}}^{\min} \leq h_{\mathbf{xc},ij}(t) \leq h_{\mathbf{xc}}^{\max}\ , \quad j \in N_{-i}. \tag{2-12}$$

The first three constraints, Equation (2-8)-(2-10), are introduced to describe the production capacity of the production unit. $p_{\mathbf{g},i}^{\min}$ and $p_{\mathbf{g},i}^{\max}$ denote the minimum and maximum electrical energy that can be produced by the micro-CHP, while $h_{\mathbf{g},i}^{\min}$ and $h_{\mathbf{g},i}^{\max}$ indicates the thermal energy produced. The minimum and maximum production energy of the boiler are denoted by $h_{\mathbf{b},i}^{\min}$ and $h_{\mathbf{b},i}^{\max}$. Similarly, Equation (2-11) refers to the capacity of the imported energy from the external party and $h_{\mathbf{im},i}^{\min}, h_{\mathbf{im},i}^{\max}$ denote the limit. The amount of energy that is exchanged between the agents at each sampling time is also limited by the capacity of the pipeline which is described by Equation (2-12) where $h_{\mathbf{xc}}^{\min} = -h_{\mathbf{xc}}^{\max}$ and $h_{\mathbf{xc}}^{\max} > 0$ defines the maximum capacity.

Up and Down Spinning Electrical Power Constraints

$$-p_{\mathbf{dg},i}(t) \leq p_{\mathbf{g},i}(t) - p_{\mathbf{d},i}(t) \leq p_{\mathbf{ug},i}(t), \tag{2-13}$$

$$p_{\mathbf{dg},i}(t) \geq 0, \quad p_{\mathbf{ug},i}(t) \geq 0, \quad \forall t \in \{k, \cdots, k+N_h-1\}, \tag{2-14}$$

where $p_{\mathbf{d},i,k}$ is the electrical energy demand of agent i at sampling time k.

Ramping Capacity Constraints

$$-p_{\mathbf{g},i}^{\mathbf{down}} \leq p_{\mathbf{g},i}(t) - p_{\mathbf{g},i}(t-1) \leq p_{\mathbf{g},i}^{\mathbf{up}}, \quad \forall t \in \{k, \cdots, k+N_h-1\}, \tag{2-15}$$

where $p_{\mathbf{g},i}^{\mathbf{down}}, p_{\mathbf{g},i}^{\mathbf{up}} > 0$ denote the maximum increase and decrease of the electrical energy produced by a micro-CHP. These constraints are only imposed on the electrical energy production due to the fact that thermal energy can be produced within each step.

Status Change Constraints

The status change constraints ensure that a micro-CHP is on (or off) for the predefined minimum value of time. $\Delta k_{\mathbf{up,g}}$ and $\Delta k_{\mathbf{down,g}}$ denote the minimum time that a micro-CHP has to be on or off. The status change constraints of a micro-CHP are as follows

$$z_{\mathbf{g},i}(t) \geq v_{\mathbf{g},i}(t) - v_{\mathbf{g},i}(t-1)\ , \quad z_{\mathbf{g},i}(t) \geq 0, \tag{2-16}$$

$$\sum_{l=t+1-\Delta k_{\mathbf{up,g}}}^{t} z_{\mathbf{g},i}(l) \leq v_{\mathbf{g},i}(t), \quad \forall t \in \{k, \cdots, k+N_h-1\}, \quad k \geq \Delta k_{\mathbf{up,g}}, \tag{2-17}$$

$$\sum_{l=t}^{t-1+\Delta k_{\mathbf{down,g}}} z_{\mathbf{g},i}(l) \leq 1 - v_{\mathbf{g},i}(t-1), \quad \forall t \in \{k, \cdots, k+N_h-1-\Delta k_{\mathbf{down,g}}\}. \tag{2-18}$$

Similarly, the status change constraints of a boiler are:

$$z_{\mathbf{b},i}(t) \geq v_{\mathbf{b},i}(t) - v_{\mathbf{b},i}(t-1)\ , \quad z_{\mathbf{b},i}(t) \geq 0, \tag{2-19}$$

$$\sum_{l=t+1-\Delta k_{\mathbf{up,b}}}^{t} z_{\mathbf{b},i}(l) \leq v_{\mathbf{b},i}(t), \quad \forall t \in \{k, \cdots, k+N_h-1\}, \quad k \geq \Delta k_{\mathbf{up,b}} \tag{2-20}$$

$$\sum_{l=t}^{t-1+\Delta k_{\mathbf{down,b}}} z_{\mathbf{b},i}(l) \leq 1 - v_{\mathbf{b},i}(t-1), \quad \forall t \in \{k, \cdots, k+N_h-1-\Delta k_{\mathbf{down,b}},\}. \tag{2-21}$$

The minimum off time ($\Delta k_{\mathbf{down}}$) is necessary to represent that a production unit require some time to shut down and start up again, that is the transient-response time of the production unit. On the other hand, the minimum on time ($\Delta k_{\mathbf{up}}$) is introduced to avoid high frequency of the switching of the production unit which could lead to inefficient fuel consumption.

Balance Constraints

By definition, the thermal energy that is exchanged between agent i and j should follow

$$h_{\mathbf{xc},ij}(t) + h_{\mathbf{xc},ji}(t) = 0, \quad \forall j \in \mathcal{N}_{-i}, \quad \forall i \in \mathcal{N}, \quad \forall t \in \{k, \cdots, k+N_h-1\}. \tag{2-22}$$

Probabilistic Constraints

One of the goals of the controller is to ensure that the thermal energy demand of each agent is fulfilled *i.e.*, the imbalance is nonnegative. Since the demand is uncertain, the constraint that ensures this condition is a chance constraint as follows

$$\mathbb{P}(x_i(t+1) \geq 0,\ \forall i \in \mathcal{N},\ \forall t \in \{k, \cdots, k+N_h-1\}\}) \geq 1-\varepsilon, \tag{2-23}$$

where $\mathbb{P}$ is the probability operator and ε is the level of violation. This implies that the imbalance error should be a nonnegative value for all the thermal energy demand realizations except for a set of probability which is at most ε.

2-2-3 Cost Function

The cost ($\boldsymbol{J}_t$) in this optimization problem is defined as a quadratic function with respect to the state and the decision variables as follows

$$\boldsymbol{J}_t = \sum_{t=k}^{k+N_h-1} \sum_{i=1}^{N} J_i(t), \tag{2-24}$$

where

$$J_i(t) = x_i(t) Q_i x_i(t) + \boldsymbol{u}_{m,i}^T(t) R_i \boldsymbol{u}_{m,i}(t), \tag{2-25}$$

in which $Q_i > 0$ and $R_i \geq 0 \quad \forall i \in \mathcal{N}$. Moreover, $R_i \in \mathbb{R}^{10\times 10}$ is a diagonal matrix, the diagonal elements of which represent the cost coefficient of the main decision variables as follows

$$\begin{aligned} R_i &= \mathrm{diag}(r_i), \\ r_i &= \begin{bmatrix} 0 & r_{\mathbf{up},i} & r_{\mathbf{dp},i} & \frac{r_{\mathbf{g},i}}{\eta_{\mathbf{CHP}}} & \frac{r_{\mathbf{b},i}}{\eta_{\mathbf{b}}} & r_{\mathbf{im},i} & 1 & 1 & 0 & 0 \end{bmatrix}, \end{aligned}$$

where $r_{\mathbf{up}}$ and $r_{\mathbf{dp},i}$ denote the cost coefficient of the up and down spinning, $r_{\mathbf{g},i}$ and $r_{\mathbf{b},i}$ denote the cost coefficient of the micro-CHP and the boiler, and $r_{\mathbf{im},i}$ denotes the cost coefficient of importing energy from a third party. Moreover, $\eta_{\mathbf{CHP}} \in (0,1)$ and $\eta_{\mathbf{b}} \in (0,1)$ denote the efficiency of the micro-CHP and the boiler. Note that the first element is zero due to the fact that $p_{\mathbf{g},i}$ and $h_{\mathbf{g},i}$ are coupled through

$$p_{\mathbf{g},i} = \frac{\eta_p}{\eta_h} h_{\mathbf{g},i}. \tag{2-26}$$

The seventh and eight entries are 1 because they correspond to the startup cost while the last two entries are zero since they belong to the auxiliary variables $z_{\mathbf{g},i}$ and $z_{\mathbf{b},i}$.

In addition, the cost coefficient $r_{\mathbf{g},i}$ and $r_{\mathbf{b},i}$ incorporate the production cost coefficient and the cost coefficient of the carbon dioxide emission as follows:

$$\begin{aligned} r_{\mathbf{g},i} &= r_{\mathbf{pg},i} + r_{\mathrm{CO}_2,\mathbf{g},i}, \\ r_{\mathbf{b},i} &= r_{\mathbf{pb},i} + r_{\mathrm{CO}_2,\mathbf{b},i}, \end{aligned}$$

where $r_{\mathbf{p}}$ denotes the cost coefficient of the production and r_{CO_2} denotes the cost coefficient of the carbon dioxide emission.

The definition of the cost function infers that we want to minimize the imbalance error as well as the efficiency. We penalize the imbalance error and the energy produced by the boiler and micro-CHP as can be seen on the definition on the weight matrices R_i. The weights assigned to the micro-CHP, the boiler, and the energy imported from the external party are set based on the economic cost of producing thermal energy. Furthermore, the energy exchanged between agents which is defined in $\boldsymbol{u}_{c,i}(k)$ is not penalized.

It is also important to emphasize that the cost function is not strictly convex with respect to all the decision variables. Not only that $R_i \ \forall i \in \mathcal{N}$ is positive semidefinite, but $\boldsymbol{u}_{c,i}(k)$ and $\boldsymbol{v}_i(k)$ for all $i \in \mathcal{N}$ and k are not taken into account in the cost. Note that the fact whether the cost is strictly convex is important when formulating a distributed optimization for the problem.

2-3 Problem Formulation

After defining the dynamics, constraints, and cost function of the system, we can write this optimization problem in a compact form. The predictions of the states, which are based on their dynamics indicated in Equation (2-5), can be written as follows

$$\tilde{\boldsymbol{x}}_i = \mathbf{A}_i x_i(k) + \mathbf{B}_i \tilde{\boldsymbol{u}}_i + \mathbf{H}_i \tilde{\boldsymbol{w}}_i, \quad \forall i \in \mathcal{N}, \tag{2-27}$$

where $\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{w}}_i, \mathbf{A}_i, \mathbf{B}_i, \mathbf{H}_i$ are given in Appendix B-1. Moreover, the local constraints which are defined in Equation (2-7) - (2-21) can also be written compactly as follows

$$\mathbf{E}_i \tilde{\boldsymbol{u}}_i + \mathbf{F}_i \tilde{\boldsymbol{v}}_i + \mathbf{P}_i \leq 0, \quad \forall i \in \mathcal{N}, \tag{2-28}$$

where $\tilde{\boldsymbol{v}}_i = [\boldsymbol{v}_i^T(k) \ \cdots \ \boldsymbol{v}_i^T(k + N_h - 1)]^T$ and $\mathbf{E}_i$, $\mathbf{F}_i$, and $\mathbf{P}_i$ are matrices with appropriate dimensions which are derived from the constraints as explained in Appendix B-2. Equation

(2-28) indicates the local constraints due to the fact that only the local variables $\boldsymbol{u}_i(t)$ and $\boldsymbol{v}_i(t) \quad \forall t \in \{k, \cdots, k+N_h-1\}$ of each agent are involved in each constraint whereas the balance constraints (Equation (2-22)) couple the decision variables of neighboring agents. Thus Equation (2-22) is called the coupling constraints and can also be written in a canonical form as

$$\tilde{\boldsymbol{u}}_{c,i} + \sum_{j \in \mathcal{N}_{-i}} \mathbf{G}_{ij} \tilde{\boldsymbol{u}}_{c,j} = 0, \quad \forall i \in \mathcal{N}, \tag{2-29}$$

where $\tilde{\boldsymbol{u}}_{c,j} = \begin{bmatrix} \boldsymbol{u}_{c,j}^T(k) & \boldsymbol{u}_{c,j}^T(k+1) & \cdots & \boldsymbol{u}_{c,j}^T(k+N_h-1) \end{bmatrix}^T$ and $\mathbf{G}_{ij} \forall j \in \mathcal{N}_{-i} \forall i \in \mathcal{N}$ are block diagonal matrices of $G_{ij} \in \mathbb{R}^{|\mathcal{N}_{-i}| \times |\mathcal{N}_{-j}|}$. In addition, $G_{ij} \forall j \in \mathcal{N}_{-i} \forall i \in \mathcal{N}$ are matrices, the elements of which are 1 and 0 such that the balance constraints are satisfied. An example of constructing matrices G_{ij} is given in Appendix B-3.

Hence, we can define the centralized optimization problem for the overall thermal grid over a finite prediction horizon N_h as

$$\begin{aligned}
\underset{\{\tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{v}}_i\}_{i=1}^N}{\text{minimize}} \quad & \sum_{i=1}^{N} \boldsymbol{J}_i(\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{u}}_i) && \text{(2-30a)} \\
\text{subject to} \quad & \mathbb{P}(\mathbf{A} x_i(k) + \mathbf{B}_i \tilde{\boldsymbol{u}}_i + \mathbf{H}_i \tilde{\boldsymbol{w}}_i \geq 0, \forall i \in \mathcal{N}) \geq 1 - \varepsilon, && \text{(2-30b)} \\
& \mathbf{E}_i \tilde{\boldsymbol{u}}_i + \mathbf{F}_i \tilde{\boldsymbol{v}}_i + \mathbf{P}_i \leq 0, && \text{(2-30c)} \\
& \tilde{\boldsymbol{u}}_{c,i} + \sum_{j \in \mathcal{N}_{-i}} \mathbf{G}_{ij} \tilde{\boldsymbol{u}}_{c,j} = 0, \quad \forall i \in \mathcal{N}, && \text{(2-30d)}
\end{aligned}$$

where

$$\boldsymbol{J}_i(\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{u}}_i) = \tilde{\boldsymbol{x}}_i^T \boldsymbol{Q}_i \tilde{\boldsymbol{x}}_i + \tilde{\boldsymbol{u}}_{m,i}^T \boldsymbol{R}_i \tilde{\boldsymbol{u}}_{m,i}, \tag{2-31}$$

in which $\boldsymbol{Q}_i$ and $\boldsymbol{R}_i$ are block diagonal matrices of Q_i and R_i as defined in Subsection 2-2-3. Furthermore, $\tilde{\boldsymbol{u}}_i^m = \begin{bmatrix} \boldsymbol{u}_{m,i}^T(k) & \boldsymbol{u}_{m,i}^T(k+1) & \cdots & \boldsymbol{u}_{m,i}^T(k+N_h-1) \end{bmatrix}^T$.

Problem (2-30) is a chance-constrained mixed-integer problem. In general, such problems are hard to solve not only due to the existence of the probabilistic constraints [3] but also the mixed-integer variables. However, in the following section, we explain a tractable method to solve this problem.

2-4 Tractable Methodology for Stochastic MPC Problems

The authors of [1] provide a tractable methodology to solve the optimization problem (2-30). It consists of two steps in which the first step is approximating the chance constraints by providing probabilistic bounds and the second step is reformulating the problem into a robust mixed-integer problem. These steps are explained in the following parts.

2-4-1 Computing the Probabilistic Bounds

The chance constraints can be approximated by constructing a bounded set of uncertainty realizations with a predefined probability of violation [18]. Since each agent has its chance constraint which correspond to its uncertain variable $w_i(k)$, we define $\mathcal{B}_i(\boldsymbol{\gamma}_i)$ to be the bounded

set of the uncertainty realizations of $w_i(k)$. We assume that $\mathcal{B}_i(\boldsymbol{\gamma}_i)$ is an axis-aligned hyper-rectangle[2] for each agent i. Then we parametrize $\mathcal{B}_i(\boldsymbol{\gamma}_i) = \times_{t=k}^{k+N_h-1} \left[\underline{\gamma}_i(t), \overline{\gamma}_i(t)\right]$ by the upper and lower bounds $\boldsymbol{\gamma}_i = (\underline{\boldsymbol{\gamma}}_i, \overline{\boldsymbol{\gamma}}_i) \in \mathbb{R}^{2N_h}$, where $\underline{\boldsymbol{\gamma}}_i = [\underline{\gamma}_i(k), \cdots, \underline{\gamma}_i(k+N_h-1)]^T \in \mathbb{R}^{N_h}$ and $\overline{\boldsymbol{\gamma}}_i = [\overline{\gamma}(k)_i, \cdots, \overline{\gamma}_i(k+N_h-1)]^T \in \mathbb{R}^{N_h}$. Now, in order to compute the bounds, we consider the following chance-constrained optimization problem:

$$\begin{aligned} \underset{\boldsymbol{\gamma}_i}{\text{minimize}} \quad & \sum_{t=k}^{k+N_h-1} \overline{\gamma}_i(t) - \underline{\gamma}_i(t) \\ \text{subject to} \quad & \mathbb{P}(w_i(t) \in \Delta_i \quad | \quad \underline{\gamma}_i(t) \leq w_i(t) \leq \overline{\gamma}_i(t), \forall t \in \{k, \cdots, k+N_h-1\}) \geq 1-\varepsilon. \end{aligned} \tag{2-32}$$

Since Problem (2-32) is convex, we can apply the scenario approach to obtain the bounds $\boldsymbol{\gamma}_i$ with a pre-determined probabilistic guarantee as follows

$$\begin{aligned} \underset{\boldsymbol{\gamma}_i}{\text{minimize}} \quad & \sum_{t=k}^{k+N_h-1} \overline{\gamma}_i(t) - \underline{\gamma}_i(t) \\ \text{subject to} \quad & \underline{\gamma}_i(t) \leq w_i^{(s)}(t) \leq \overline{\gamma}_i(t) \quad \begin{cases} \forall t \in \{k, \cdots, k+N_h-1\}, \\ \forall s \in \{1, \cdots, N_s\}, \end{cases} \end{aligned} \tag{2-33}$$

where $w_i^{(s)}$ denote the s-th scenario of w_i and the number of scenarios N_s is determined by Equation (2-34) as the following [3]:

$$N_s \geq \frac{2}{\varepsilon}\left(N_d + \ln\frac{1}{\beta}\right), \tag{2-34}$$

where $N_d = 2NN_h$. Thus the optimal solution $(\boldsymbol{\gamma}_i^\star)$ that is obtained by solving Problem (2-33) is a feasible solution for Problem (2-32) with confidence level of $1-\beta$. It is important to note that we only need a finite number of instances of $w_i(k)$, and it is not required that the probability space Δ_i and the probability measure $\mathbb{P}$ to be known explicitly [1].

2-4-2 Robust Reformulation

After $\mathcal{B}_i(\boldsymbol{\gamma}_i^\star)$ $\forall i \in \mathcal{N}$ are constructed, we can formulate the robust counterpart of Problem (2-30) where $\mathbf{W}_i(k) \in \mathcal{B}_i \cap \Delta_i \forall i \in \mathcal{N}$. As stated in [1], it is a finite-horizon mixed-integer problem any feasible solution of which is also a feasible solution of Problem (2-30) with confidence level of $1-\beta$. The robust counterpart is

$$\underset{\{\tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{v}}_i\}_{i=1}^N}{\text{minimize}} \quad \sum_{i=1}^{N} \boldsymbol{J}_i(\tilde{\boldsymbol{u}}_i) \tag{2-35a}$$

$$\text{subject to} \quad \mathbf{A}_i x_i(k) + \mathbf{B}_i \tilde{\boldsymbol{u}}_i + \mathbf{H}_i(\boldsymbol{\gamma}_i^o + \boldsymbol{\eta}_i) \geq 0, \tag{2-35b}$$

$$\mathbf{E}_i \tilde{\boldsymbol{u}}_i + \mathbf{F}_i \tilde{\boldsymbol{v}}_i + \mathbf{P}_i \leq 0, \tag{2-35c}$$

$$\tilde{\boldsymbol{u}}_{c,i} + \sum_{j \in \mathcal{N}_{-i}} \mathbf{G}_{ij} \tilde{\boldsymbol{u}}_{c,j} = 0, \qquad \forall i \in \mathcal{N}, \tag{2-35d}$$

[2]The choice of a hyper-rectangle is not restrictive and any convex set with convex volume could have been chosen instead [18]

where $\boldsymbol{\gamma_i^o} = [\gamma_i^o(k), \cdots, \gamma_i^o(k+N_h-1)]^T \in \mathbb{R}^{N_h}, \quad \forall i \in \mathcal{N}$ are vectors, elements of which are the middle points of the hyper-rectangle $\mathcal{B}_i$, *i.e.,* defined as $\boldsymbol{\gamma_i^o} = 0.5(\overline{\gamma}_i^\star + \underline{\gamma}_i^\star)$. Moreover, $\boldsymbol{\eta_i} = [\eta_i(k) \ \cdots \ \eta_i(k+N_h-1)]^T \forall i \in \mathcal{N}$ are the bounds for the worst-case realizations of $w_i(t) \ \forall t \in \{k, \cdots, k+N_h-1\}$, that is

$$\begin{aligned} \underset{\boldsymbol{\eta_i}}{\text{maximize}} \quad & \boldsymbol{\eta_i} \\ \text{subject to} \quad & \eta_i(t) \leq \overline{\gamma}_i^\star(t) - \gamma_i^o(t), \\ & \eta_i(t) \leq \underline{\gamma}_i^\star(t) - \gamma_i^o(t), \quad \forall t \in \{k, \cdots, k+N_h-1\}. \end{aligned} \tag{2-36}$$

Note that since now we have

$$\tilde{\boldsymbol{x}}_i = \mathbf{A}_i x_i(k) + \mathbf{B}_i \tilde{\boldsymbol{u}}_i + \mathbf{H}_i(\boldsymbol{\gamma_i^o} + \boldsymbol{\eta_i}),$$

$\tilde{\boldsymbol{x}}_i$ can be substituted so that we obtain an expression of $\boldsymbol{J}_i$ only as a function of $\tilde{\boldsymbol{u}}_i$.

The stochastic procedure which is called as the robust randomized MPC is stated in the following Algorithm 1.

Algorithm 1 Robust Randomized MPC

for $k = 1, 2, \ldots$ **do**
- Set $\varepsilon \in (0,1), \ \beta \in (0,1)$.
- Generate N_s scenarios where N_s is computed as in Equation (2-34).
- Establish $\mathcal{B}_i(\boldsymbol{\gamma}_i^\star) \forall i \in \mathcal{N}$ by solving Problem (2-32).
- Solve the robust problem (2-35)
- Apply $\boldsymbol{u}_i(k) \forall i \in \mathcal{N}$ to the agents.

end for

2-5 Distributed Robust Randomized MPC

One of the main contribution of this thesis is extending the robust randomized MPC to a distributed approach. This is obtained by decomposing the proposed robust reformulation, which is defined in (2-35). Due to the existence of the coupling constraints (Equation (2-35d)), it is not a trivially separable problem. Thus we will decompose it using a dual decomposition method. It is also known that since $R_i \geq 0 \ \forall i \in \mathcal{N}$, the cost function is not strictly convex. This means that the standard dual decomposition method will not work [27]. Hence ADMM is proposed to be applied to decompose and solve it in a distributed setting. In this section, two ADMM formulations for this problem are explained, one is a fully distributed scheme while the other is with a coordinator.

2-5-1 Fully Distributed Coordination based on MPC

Problem (2-35) can be expressed in a compact form as

$$\begin{aligned} \underset{\{\tilde{\boldsymbol{u}}_i,\tilde{\boldsymbol{v}}_i\}_{i=1}^N}{\text{minimize}} \quad & \sum_{i=1}^{N} \boldsymbol{J}_i(\tilde{\boldsymbol{u}}_i) && \text{(2-37a)}\\ \text{subject to} \quad & \tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{v}}_i \in \mathcal{L}_i, && \text{(2-37b)}\\ & \tilde{\boldsymbol{u}}_{c,i} \in \mathcal{C}_i, \quad \forall i \in \mathcal{N}, && \text{(2-37c)} \end{aligned}$$

where $\mathcal{L}_i$ is the set which is defined by the local constraints (2-35b) and (2-35c) of agent i while $\mathcal{C}_i$ is the set which is defined by the coupling constraint (2-35d). We can reformulate the problem into

$$\begin{aligned} \underset{\{\tilde{\boldsymbol{u}}_i,\tilde{\boldsymbol{v}}_i\}_{i=1}^N}{\text{minimize}} \quad & \sum_{i=1}^{N} \boldsymbol{J}_i(\tilde{\boldsymbol{u}}_i) + \boldsymbol{\psi}_i(\tilde{\boldsymbol{y}}_i) && \text{(2-38a)}\\ \text{subject to} \quad & \tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{v}}_i \in \mathcal{L}_i, && \text{(2-38b)}\\ & \tilde{\boldsymbol{u}}_{c,i} - \tilde{\boldsymbol{y}}_i = 0 \quad \forall i \in \mathcal{N}, && \text{(2-38c)} \end{aligned}$$

in which $\boldsymbol{\psi}_i(\tilde{\boldsymbol{y}}_i)$ is the indicator function of $\mathcal{C}_i$, that is $\psi_i(\tilde{\boldsymbol{y}}_i) = 0$ if $\tilde{\boldsymbol{y}}_i \in \mathcal{C}_i$ and $\psi_i(\tilde{\boldsymbol{y}}_i) = +\infty$ otherwise. Note that $\tilde{\boldsymbol{y}}_i = [\boldsymbol{y}_i(k) \ \cdots \ \boldsymbol{y}_i(k+N_h-1)] \in \mathbb{R}^{N_h|\mathcal{N}_{-i}|}$, $\forall i \in \mathcal{N}$. By introducing the Lagrange multipliers, $\tilde{\boldsymbol{\lambda}}_i = [\boldsymbol{\lambda}_i^T(k) \ \cdots \boldsymbol{\lambda}_i^T(k+N_h-1)]^T \ \forall i \in \mathcal{N}$, where $\boldsymbol{\lambda}_i(t) \in \mathbb{R}^{|\mathcal{N}_{-i}|} \ \forall t \in \{k, \cdots, k+N_h-1\}$ and $\rho > 0$ as the penalty parameter, the augmented Lagrangian of this problem can be written as

$$\begin{aligned} \boldsymbol{L}_\rho &= \sum_{i=1}^{N} \left(\boldsymbol{J}_i(\tilde{\boldsymbol{u}}_i) + \boldsymbol{\psi}_i(\tilde{\boldsymbol{y}}_i) + \tilde{\boldsymbol{\lambda}}_i^T(\tilde{\boldsymbol{u}}_{c,i} - \tilde{\boldsymbol{y}}_i) + \frac{\rho}{2}\|\tilde{\boldsymbol{u}}_{c,i} - \tilde{\boldsymbol{y}}_i\|_2^2 \right) \\ &= \sum_{i=1}^{N} \boldsymbol{L}_{\rho,i}(\tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{y}}_i, \tilde{\boldsymbol{\lambda}}_i). \end{aligned} \tag{2-39}$$

Hence, the augmented dual problem associated with Problem (2-35) is

$$\begin{aligned} \underset{\{\tilde{\boldsymbol{u}}_i,\tilde{\boldsymbol{v}}_i\}_{i=1}^N}{\text{minimize}} \quad & \sum_{i=1}^{N} \boldsymbol{L}_{\rho,i}(\tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{y}}_i, \tilde{\boldsymbol{\lambda}}_i) && \text{(2-40a)}\\ \text{subject to} \quad & \tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{v}}_i \in \mathcal{L}_i, \quad \forall i \in \mathcal{N}, && \text{(2-40b)} \end{aligned}$$

which is separable. Thus the ADMM that solves this problem consists of the iterations of:

1. Updating $\tilde{\boldsymbol{u}}_i$ and $\tilde{\boldsymbol{v}}_i$ for all $i \in \mathcal{N}$:

$$\begin{aligned} \tilde{\boldsymbol{u}}_i^{(q+1)}, \tilde{\boldsymbol{v}}_i^{(q+1)} := \underset{\tilde{\boldsymbol{u}}_i,\tilde{\boldsymbol{v}}_i}{\text{argmin}} \quad & \boldsymbol{L}_{\rho,i}(\tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{y}}_i^{(q)}, \tilde{\boldsymbol{\lambda}}_i^{(q)}) \\ \text{subject to} \quad & \tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{v}}_i \in \mathcal{L}_i. \end{aligned} \tag{2-41}$$

2. Updating $\tilde{\boldsymbol{y}}_i$ for all $i \in \mathcal{N}$:

$$\tilde{\boldsymbol{y}}_i^{(q+1)} = \Pi_{\mathcal{C}_i}\left(\tilde{\boldsymbol{u}}_{c,i}^{(q+1)}\right). \tag{2-42}$$

$\Pi_{\mathcal{C}_i}(\cdot)$ is the projection onto the set $\mathcal{C}_i$ which is

$$\begin{aligned} \tilde{\boldsymbol{y}}_i = \underset{\tilde{\boldsymbol{y}}_i}{\text{argmin}} \quad & ||\tilde{\boldsymbol{y}}_i - \tilde{\boldsymbol{u}}_{c,i}^{(q+1)}||_2 \\ \text{subject to } & \tilde{\boldsymbol{y}}_i + \sum_{j \in \mathcal{N}_{-i}} \mathbf{G}_{ij} \tilde{\boldsymbol{u}}_{c,j}^{(q+1)} = 0. \end{aligned} \tag{2-43}$$

Clearly, the solution of this minimization problem is

$$\tilde{\boldsymbol{y}}_i^{(q+1)} = - \sum_{j \in \mathcal{N}_{-i}} \mathbf{G}_{ij} \tilde{\boldsymbol{u}}_{c,j}^{(q+1)}. \tag{2-44}$$

Thus we use Equation (2-44) as the update rule of $\tilde{\boldsymbol{y}}_i$.

3. Updating $\tilde{\boldsymbol{\lambda}}_i$ for all $i \in \mathcal{N}$ via a gradient method:

$$\tilde{\boldsymbol{\lambda}}_i^{(q+1)} = \tilde{\boldsymbol{\lambda}}_i^{(q)} + \rho \left(\tilde{\boldsymbol{u}}_{c,i}^{(q+1)} - \tilde{\boldsymbol{y}}_i^{(q+1)} \right). \tag{2-45}$$

The algorithm stops when $\left(\tilde{\boldsymbol{u}}_{c,i}^{(q)} - \tilde{\boldsymbol{y}}_i^{(q)} \right) \rightarrow 0, \forall i \in \mathcal{N}$, thus the stopping criterion of the algorithm is given by

$$\left|\left| \begin{bmatrix} \tilde{\boldsymbol{u}}_{c,1}^{(q)} - \tilde{\boldsymbol{y}}_1^{(q)} \\ \vdots \\ \tilde{\boldsymbol{u}}_{c,N}^{(q)} - \tilde{\boldsymbol{y}}_N^{(q)} \end{bmatrix} \right|\right|_2 < \epsilon, \tag{2-46}$$

for a small $\epsilon > 0$.

The distributed version of the robust randomized MPC is written in Algorithm 2. Note that the steps of updating $\tilde{\boldsymbol{y}}_i$ and $\tilde{\boldsymbol{\lambda}}_i$ is fully distributed to the agents. Even so, it is also possible to let a coordinator does these steps. Moreover, not all decision variables but only $\tilde{\boldsymbol{u}}_{c,i}$ which contains $h_{\mathbf{xc},ij}$ that needs to be communicated between the agents (in a fully distributed setting) or to the coordinator.

One issue of an iterative algorithm such as this method is that it possibly requires a large number of iterations before the stopping criteria is met. In this regard, warm start is a method which could reduce the number of iterations. It is done by using the solutions obtained in the previous sampling time since it often gives a good enough approximation to result in far fewer iterations [27]. For instance, let say that $\tilde{\boldsymbol{y}}_i^{(q)}$ and $\tilde{\boldsymbol{\lambda}}_i^{(q)}$ are the solutions obtained at the last iteration (q) at sampling time k, thus we initialize $\tilde{\boldsymbol{y}}_i^{(0)}$ and $\tilde{\boldsymbol{\lambda}}_i^{(0)}$ for the next sampling time $(k+1)$ as

$$\tilde{\boldsymbol{y}}_i^{(0)} = \begin{bmatrix} \tilde{\boldsymbol{y}}_i^{(q)T}(k+1) & \cdots & \tilde{\boldsymbol{y}}_i^{(q)T}(k+N_h-1) & \mathbf{0} \end{bmatrix}^T, \tag{2-47}$$

$$\tilde{\boldsymbol{\lambda}}_i^{(0)} = \begin{bmatrix} \tilde{\boldsymbol{\lambda}}_i^{(q)T}(k+1) & \cdots & \tilde{\boldsymbol{\lambda}}_i^{(q)T}(k+N_h-1) & \mathbf{0} \end{bmatrix}^T. \tag{2-48}$$

Algorithm 2 Fully Distributed Robust Randomized MPC

Initialize $\tilde{\boldsymbol{y}}_i^{(0)},\ \tilde{\boldsymbol{\lambda}}_i^{(0)},\ \forall i \in \mathcal{N},\ \rho > 0$.
for $k = 1, 2, \ldots$ **do**
Set $\varepsilon \in (0,1),\ \beta \in (0,1)$.
Generate N_s scenarios where N_s is computed as in Equation (2-34).
Establish $\mathcal{B}_i(\boldsymbol{\gamma}_i^\star) \forall i \in \mathcal{N}$ by solving Problem (2-32).
$q = 1$.
while $\left\| \begin{bmatrix} \tilde{\boldsymbol{u}}_{c,1}^{(q-1)} - \tilde{\boldsymbol{y}}_1^{(q-1)} \\ \vdots \\ \tilde{\boldsymbol{u}}_{c,N}^{(q-1)} - \tilde{\boldsymbol{y}}_N^{(q-1)} \end{bmatrix} \right\|_2 \geq \epsilon$ **do**
for $i \in \mathcal{N}$ **do**
Each agent solves the local optimization to obtain $\tilde{\boldsymbol{u}}_i^{(q)}$ and $\tilde{\boldsymbol{v}}_i^{(q)}$, that is

$$\begin{aligned} \underset{\tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{v}}_i}{\text{minimize}} \quad & L_{\rho,i}\left(\tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{y}}_i^{(q-1)}, \tilde{\boldsymbol{\lambda}}_i^{(q-1)}\right) \\ \text{subject to} \quad & \text{(2-49b) and (2-49c).} \end{aligned}$$

end for
Exchange $\tilde{\boldsymbol{u}}_{c,i}^{(q)}$ with the neighbors.
for $i \in \mathcal{N}$ **do**
Update $\tilde{\boldsymbol{y}}_i^{(q)}$:

$$\tilde{\boldsymbol{y}}_i^{(q)} = -\sum_{j \in \mathcal{N}_{-i}} \mathbf{G}_{ij} \tilde{\boldsymbol{u}}_{c,j}^{(q)}.$$

Update $\tilde{\boldsymbol{\lambda}}_i^{(q)}$:

$$\tilde{\boldsymbol{\lambda}}_i^{(q)} = \tilde{\boldsymbol{\lambda}}_i^{(q-1)} + \rho\left(\tilde{\boldsymbol{u}}_{c,i}^{(q)} - \tilde{\boldsymbol{y}}_i^{(q)}\right).$$

end for
$q = q + 1$.
end while
Apply $\boldsymbol{u}_i^{(q)}(k)\ \forall i \in \mathcal{N}$.
Warm start: Initialize $\tilde{\boldsymbol{y}}_i^{(0)},\ \tilde{\boldsymbol{\lambda}}_i^{(0)},\ \forall i \in \mathcal{N}$, for the next sampling time with a warm start using $\tilde{\boldsymbol{y}}_i^{(q)},\ \tilde{\boldsymbol{\lambda}}_i^{(q)}$ from the last iteration:

$$\begin{aligned} \tilde{\boldsymbol{y}}_i^{(0)} &= \begin{bmatrix} \tilde{\boldsymbol{y}}_i^{(q)T}(k+1) & \cdots & \tilde{\boldsymbol{y}}_i^{(q)T}(k+N_h-1) & \mathbf{0} \end{bmatrix}^T, \\ \tilde{\boldsymbol{\lambda}}_i^{(0)} &= \begin{bmatrix} \tilde{\boldsymbol{\lambda}}_i^{(q)T}(k+1) & \cdots & \tilde{\boldsymbol{\lambda}}_i^{(q)T}(k+N_h-1) & \mathbf{0} \end{bmatrix}^T. \end{aligned}$$

end for

2-5-2 Distributed ADMM with Coordinator

The second ADMM method is formulated by perceiving the problem as an optimal exchange problem which is discussed in Appendix C-2-2. In this regard, we can consider the robust reformulation of the problem as follows:

$$\begin{aligned}
\underset{\{\tilde{\boldsymbol{u}}_i,\tilde{\boldsymbol{v}}_i\}_{i=1}^N}{\text{minimize}} \quad & \sum_{i=1}^{N} \boldsymbol{J}_i(\tilde{\boldsymbol{u}}_i) && \text{(2-49a)}\\
\text{subject to} \quad & \mathbf{A}_i x_i(k) + \mathbf{B}_i\tilde{\boldsymbol{u}}_i + \mathbf{H}_i(\boldsymbol{\gamma}_i^{\boldsymbol{o}} + \boldsymbol{\eta}_i) \geq 0, && \text{(2-49b)}\\
& \mathbf{E}_i\tilde{\boldsymbol{u}}_i + \mathbf{F}_i\tilde{\boldsymbol{v}}_i + \mathbf{P}_i \leq 0, \quad \forall i \in \mathcal{N}, && \text{(2-49c)}\\
& \sum_{i=1}^{N} \mathbf{K}_i\tilde{\boldsymbol{u}}_{c,i} = 0, && \text{(2-49d)}
\end{aligned}$$

in which the formulation of the coupling constraint is slightly modified. Instead of having N coupling constraints as in Problem (2-35), in this formulation there is only one global coupling constraints which is Equation (2-49d). Moreover, $\mathbf{K}_i\ \forall i \in \mathcal{N}$ are diagonal matrices of K_i, the elements of which are 0 and 1 such that the balance constraints are satisfied (see Appendix B-3). Hence, we can follow the unscaled form of ADMM for such problems as provided in [27], in order to solve it. This consists following iterations:

1. Updating of $\tilde{\boldsymbol{u}}_i$ and $\tilde{\boldsymbol{v}}_i$:

$$\begin{aligned}
\tilde{\boldsymbol{u}}_i^{(q+1)}, \tilde{\boldsymbol{v}}_i^{(q+1)} = \underset{\tilde{\boldsymbol{u}}_i,\tilde{\boldsymbol{v}}_i}{\text{argmin}} \quad & \boldsymbol{J}_i(\tilde{\boldsymbol{u}}_i) + \boldsymbol{\lambda}^{(q)T}\left(\mathbf{K}_i\tilde{\boldsymbol{u}}_{c,i}\right) + \frac{\rho}{2}\left|\left|\mathbf{K}_i\tilde{\boldsymbol{u}}_{c,i} + \left(\tilde{\boldsymbol{y}}^{(q)} - \mathbf{K}_i\tilde{\boldsymbol{u}}_{c,i}^{(q)}\right)\right|\right|_2^2\\
\text{subject to} \quad & \mathbf{A}_i x_i(k) + \mathbf{B}_i\tilde{\boldsymbol{u}}_i + \mathbf{H}_i(\boldsymbol{\gamma}_i^{\boldsymbol{o}} + \boldsymbol{\eta}_i) \geq 0,\\
& \mathbf{E}_i\tilde{\boldsymbol{u}}_i + \mathbf{F}_i\tilde{\boldsymbol{v}}_i + \mathbf{P}_i \leq 0.
\end{aligned} \tag{2-50}$$

2. Updating of $\tilde{\boldsymbol{y}}$:

$$\tilde{\boldsymbol{y}}^{(q+1)} = \sum_{i=1}^{N} \mathbf{K}_i\tilde{\boldsymbol{u}}_{c,i}^{(q+1)}. \tag{2-51}$$

3. Updating of $\tilde{\boldsymbol{\lambda}}$:

$$\tilde{\boldsymbol{\lambda}}^{(q+1)} = \tilde{\boldsymbol{\lambda}}^{(q)} + \rho\tilde{\boldsymbol{y}}^{(q+1)}. \tag{2-52}$$

In this approach, beside the optimization problem is solved in a distributed fashion, the process of updating the auxiliary decision variable $\tilde{\boldsymbol{y}}$ and the Lagrange multiplier $\tilde{\boldsymbol{\lambda}}$ requires a coordinator which receives the decision of $\tilde{\boldsymbol{u}}_{c,i}$ at each iteration from all agents. Algorithm 3 provides the summary of the robust randomized method in this setting.

2-6 Numerical Study of the First Case

In this section, we present a simulation study to illustrate the performance of the proposed distributed algorithms which would be compared with the performance of the centralized version of this algorithm. All simulations were carried out in MATLAB. We used YALMIP [38] and Gurobi [39] to solve the local optimization problems which are mixed-integer.

Algorithm 3 Distributed Coordinated Robust Randomized MPC

Initialize $\tilde{\boldsymbol{y}}_i^{(0)},\ \tilde{\boldsymbol{\lambda}}_i^{(0)},\ \forall i \in \mathcal{N},\ \rho > 0$.
for $k = 1, 2, \ldots$ **do**
Set $\varepsilon \in (0,1),\ \beta \in (0,1)$.
Generate N_s scenarios where N_s is computed as in Equation (2-34).
Establish $\mathcal{B}_i(\boldsymbol{\gamma}_i^\star) \forall i \in \mathcal{N}$ by solving Problem (2-32).
$q = 1$.
while $\left\| \begin{bmatrix} \tilde{\boldsymbol{u}}_{c,1}^{(q-1)} - \tilde{\boldsymbol{y}}_1^{(q-1)} \\ \vdots \\ \tilde{\boldsymbol{u}}_{c,N}^{(q-1)} - \tilde{\boldsymbol{y}}_N^{(q-1)} \end{bmatrix} \right\|_2 \geq \epsilon$ **do**
for $i \in \mathcal{N}$ **do**
Each agent solves the local optimization to obtain $\tilde{\boldsymbol{u}}_i^{(q)}$ and $\tilde{\boldsymbol{v}}_i^{(q)}$, that is

$$\begin{aligned} \underset{\tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{v}}_i}{\text{minimize}} \quad & L_{\rho,i}\left(\tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{y}}_i^{(q-1)}, \tilde{\boldsymbol{\lambda}}_i^{(q-1)}\right) \\ \text{subject to} \quad & \text{(2-35b) and (2-35c).} \end{aligned}$$

end for
All agents send their decision of $\tilde{\boldsymbol{u}}_{c,i}^{(q)}$ to the coordinator.
The coordinator updates $\tilde{\boldsymbol{y}}$ and $\tilde{\boldsymbol{\lambda}}$ as follows:

$$\tilde{\boldsymbol{y}}^{(q)} = \sum_{i=1}^{N} \mathbf{K}_i \tilde{\boldsymbol{u}}_{c,i}^{(q)}, \tag{2-53}$$

$$\tilde{\boldsymbol{\lambda}}^{(q)} = \tilde{\boldsymbol{\lambda}}^{(q-1)} + \rho \tilde{\boldsymbol{y}}^{(q)}. \tag{2-54}$$

$q = q + 1$.
end while
All agents apply $\boldsymbol{u}_i^{(q)}(k)\ \forall i \in \mathcal{N}$.
Warm start: The coordinator initializes $\tilde{\boldsymbol{y}}_i^{(0)},\ \tilde{\boldsymbol{\lambda}}_i^{(0)},\ \forall i \in \mathcal{N}$, for the next sampling time with a warm start using $\tilde{\boldsymbol{y}}_i^{(q)},\ \tilde{\boldsymbol{\lambda}}_i^{(q)}$ from the last iteration:

$$\begin{aligned} \tilde{\boldsymbol{y}}_i^{(0)} &= \begin{bmatrix} \tilde{\boldsymbol{y}}_i^{(q)T}(k+1) & \cdots & \tilde{\boldsymbol{y}}_i^{(q)T}(k+N_h-1) & \mathbf{0} \end{bmatrix}^T, \\ \tilde{\boldsymbol{\lambda}}_i^{(0)} &= \begin{bmatrix} \tilde{\boldsymbol{\lambda}}_i^{(q)T}(k+1) & \cdots & \tilde{\boldsymbol{\lambda}}_i^{(q)T}(k+N_h-1) & \mathbf{0} \end{bmatrix}^T. \end{aligned}$$

and assigns them to the corresponding agents.
end for

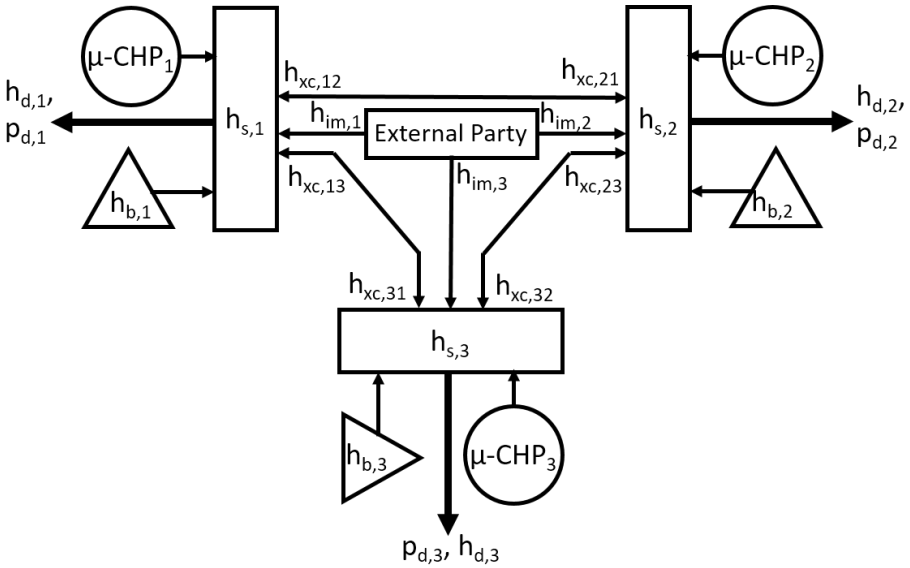


Figure 2-2: Diagram of the thermal grid considered in the simulation. It consists of three agents, each of which has a boiler, a micro-CHP, and a storage. It is also possible to exchange energy between them.

2-6-1 Simulation Setup

A thermal grid which consists of three agents ($N = 3$) and has uncertain thermal energy demand was considered as an example for the simulation study of this case. Each agent has its local thermal and electrical energy demand. They also had a micro-CHP and a boiler as the production units. Additionally, they were able to exchange energy which are denoted by $h_{\mathbf{xc},ij}\ \forall i \in \mathcal{N}, \forall j \in \mathcal{N}_{-i}$ and $|\mathcal{N}_{-i}| = 2$ for all agents. Furthermore, there was an external party which exports thermal and electrical energy to the agents if necessary. Figure 2-2 depicts the diagram of the thermal grid.

We solved the production planning problem of this system for one day in a centralized setting using Algorithm 1 and distributed settings using Algorithm 2 and 3 with $N_h = 24$. It was assumed that the operating points that were set by solving this problem at each sampling time could always be achieved by the production units. The thermal energy demand profiles of all agent for the day are shown in Figure 2-3. The parameters of each agent, such as the efficiency and the limit of the production units, the cost coefficients, are given in Table 2-1. Moreover, we assumed that the up and down ramping capacity of the electrical power were $p^{\mathbf{up}}_{\mathbf{g},i} = p^{\mathbf{down}}_{\mathbf{g},i} = p^{\mathbf{max}}_{\mathbf{g},i}/3$. Furthermore, the level of violation and confidence of the robust randomized MPC method were set to be 0.1 and 0.0001, respectively. The N_s scenarios were generated by the Markov-chain based model which is explained in Section 2-1. In order to have a clear performance comparison of the distributed and centralized methods, the probabilistic bounds were computed offline. Hence at each sampling time, all algorithms used the same bounds when solving the robust problem.

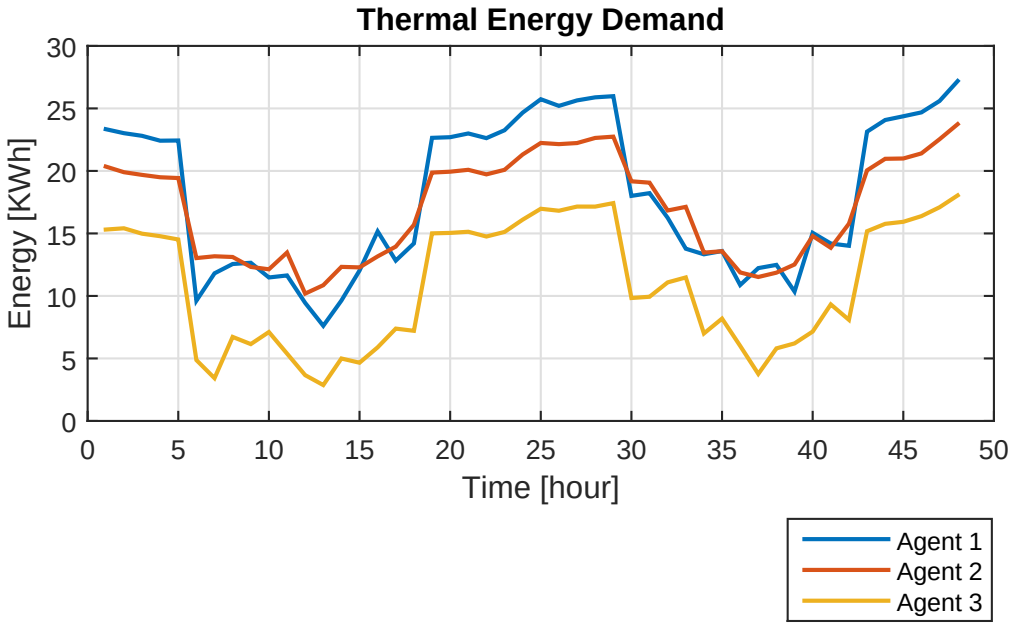


Figure 2-3: The thermal energy demand of each agent.

Table 2-1: The Values of the Parameters

Parameter	Agent 1	Agent 2	Agent 3	Unit
$p_{\mathbf{g}}^{\max}$,$h_{\mathbf{g}}^{\max}$,$p_{\mathbf{g}}^{\min}$,$h_{\mathbf{g}}^{\min}$	15, 15, 0, 2	15, 15, 0, 2	15, 15, 0, 2	KWh
$h_{\mathbf{b}}^{\max}$,$h_{\mathbf{im}}^{\max}$,$h_{\mathbf{b}}^{\min}$,$h_{\mathbf{im}}^{\min}$	30, 120, 2, 0	30, 120, 2, 0	30, 120, 2, 0	KWh
$h_{\mathbf{xc}}^{\max}$,$h_{\mathbf{xc}}^{\min}$,$x[0]$	2, -2, 2	2, -2, 1	2, -2, 0	KWh
$\eta_{\mathbf{CHP}}$,$\eta_{\mathbf{h}}$,$\eta_{\mathbf{p}}$	0.85, 0.7, 0.3	0.85, 0.7, 0.3	0.85, 0.7, 0.3	-
$\eta_{\mathbf{b}}$,$\eta_{\mathbf{s}}$,α_{ij}	0.8, 0.9, 0.25	0.8, 0.9, 0.25	0.8, 0.9, 0.25	-
$c_{\mathbf{g}}$,$c_{\mathbf{up}}$,$c_{\mathbf{dp}}$,$c_{\mathbf{b}}$	16, 0, 121, 36	25, 0, 144, 64	25, 0, 100, 400	-
$c_{\mathbf{im}}$, $c_{\mathbf{g}}^{\mathrm{su}}$,$c_{\mathbf{b}}^{\mathrm{su}}$, Q	2500, 1, 1, 100	2500, 1, 1, 100	2500, 1, 1, 100	-
Δt_{up}, Δt_{down}	2, 2	2, 2	2, 2	hours

2-6-2 Discussion of the Simulation Results

Figure 2-4 until 2-10 depict the simulation results of the three algorithms that were implemented to the system. The thermal demands of all agents were fulfilled at every sampling time and this is shown by the positive values of the imbalance errors (Figure 2-4). In order to achieve this, not all production units were required to be on (Figure 2-8 - 2-9). The boiler of Agent 3 was off all the time.

The decision that the boiler of Agent 3 was always off was logical since we set the cost coefficient to be the highest ($c_{\mathbf{b},3} = 400$) among all production units which made it as the last option in order to achieve an efficient production. Since the rest of the production units along with the possibility of exchanging energy were able to satisfy all demands, it was not necessary to use the boiler of Agent 3. By comparing Figure 2-5 and 2-6, we can see that the micro-CHPs were asked to produce more energy than the boilers. This was also due to the fact that the cost coefficients of the micro-CHPs were relatively lower than the ones of the boilers.

We can see that the balance constraints of the exchanged energy were satisfied in Figure 2-7. As mentioned earlier, it also showed that the controller took the advantage of having possibility of exchanging energy. It is seen in the figure that Agent 1 exported some of its energy to Agent 2 and 3. This resulted in the reducing of the amount of energy that had to be produced by those agents. Moreover, during 6 until 18 hours, Agent 3 also exported some of the energy produced to Agent 2.

From all figures, we can see that the distributed algorithms found approximately the same solution at every sampling time as the centralized algorithm. Figure 2-10 shows that the computed cost of the solutions of the distributed and centralized algorithm were almost identical at each sampling time. The total cost of the fully distributed and the coordinated distributed algorithm were only approximately 1% higher than the total cost of the centralized one. Therefore, it can be inferred that the proposed distributed algorithms were able to keep the performance of the centralized version while having the advantages of being distributed methods.

Theoretically, the solutions of the distributed and centralized algorithm do not necessarily have to be exactly equal due to the randomization method that was employed. In fact, the solutions of either method from one and another simulation can be different. However, in this study the bounds that were assigned for solving the robust program via centralized and distributed methods were the same.

Figure 2-11 shows the number of iterations and the computational time required by the distributed approaches to solve the problem at each sampling time. It can be seen that both distributed methods had similar number of iterations as well as computational times. However, in the simulation of the last hour, the distributed algorithm with coordinator failed to converge (satisfy the stopping criterion) in a given time and the algorithm stopped prematurely. Moreover, the computational time at each sampling hour was also considered to be too long. In half of the sampling times (from 6 until 18 hours), the distributed approaches required almost one hour to solve the problem. The simulations of both methods were done in parallel at the same time with the same computer and the local optimizations were done sequentially. Even though these two reasons affected the high computational time and it is

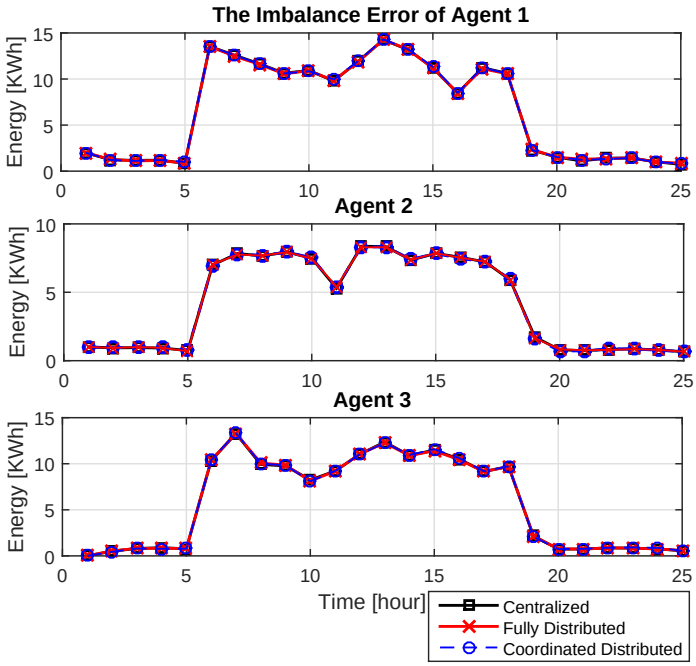


Figure 2-4: The imbalance of each agent.

believed that if the local optimizations were done in parallel we will be able to reduce the computational time, the algorithms still have a room of improvement to reduce the convergence rate.

Related to this, a difficulty that was faced when using the proposed distributed approaches comes from the choice of the penalty term ρ which is also the step size of the ascent method. In this study, it was considered as a tuning parameter and in the simulations both algorithms used the same penalty term (step size $\rho = 2$) and the same stopping criterion. The observation that can be concluded is that if ρ is too small, the convergence rate will be very slow. On the other hand, the solution will not converge if ρ is too large. From this figure, it is also noticed that the largest number of iterations happened at the first sampling time in which the Lagrange multipliers were initialized randomly. So in other words, it also shows that warm start that was implemented in both algorithms was able to reduce the number of iterations.

In summary, the simulation results show that the distributed approaches did not decrease the performance of the robust randomized method in term of the cost. Nevertheless, it is important to note that these approaches require the agents to be cooperative. Furthermore, choosing the suitable penalty term (step size) and in general increasing the convergence rate are the challenges which are important to improve these methods.

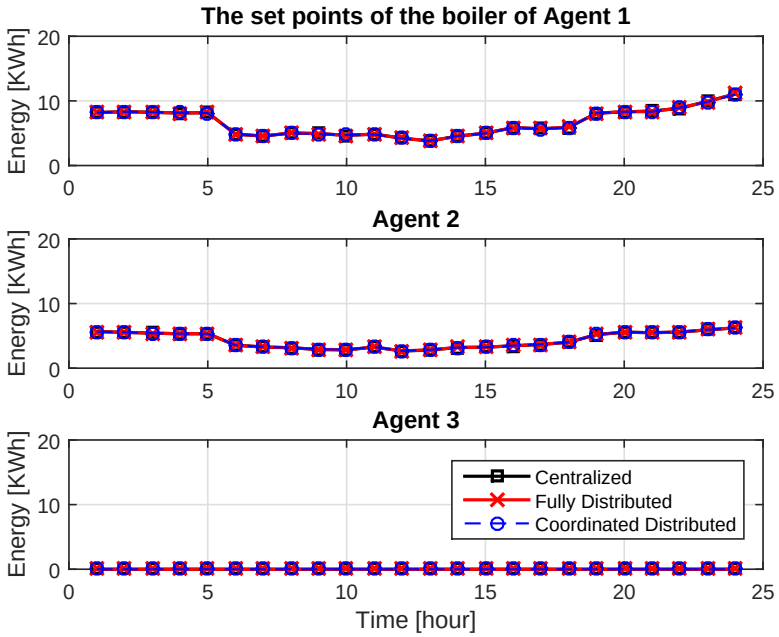


Figure 2-5: The energy production of the boilers.

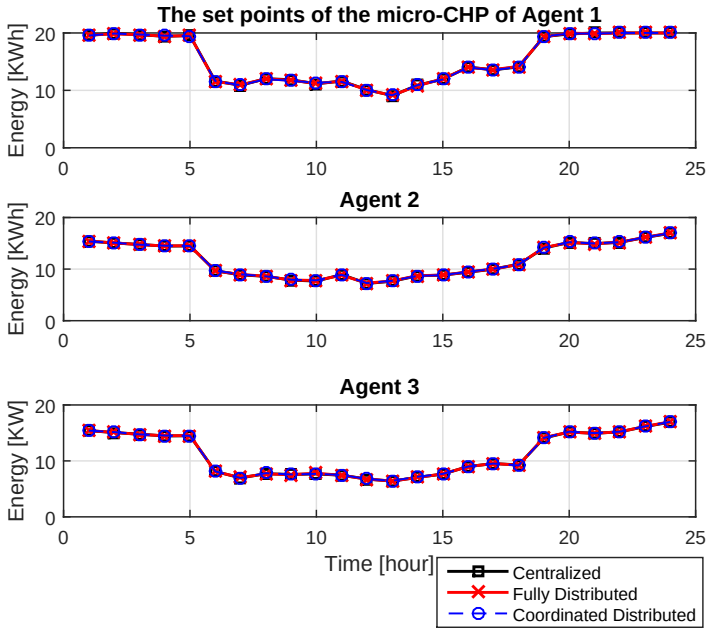


Figure 2-6: The energy production of the micro-CHPs.

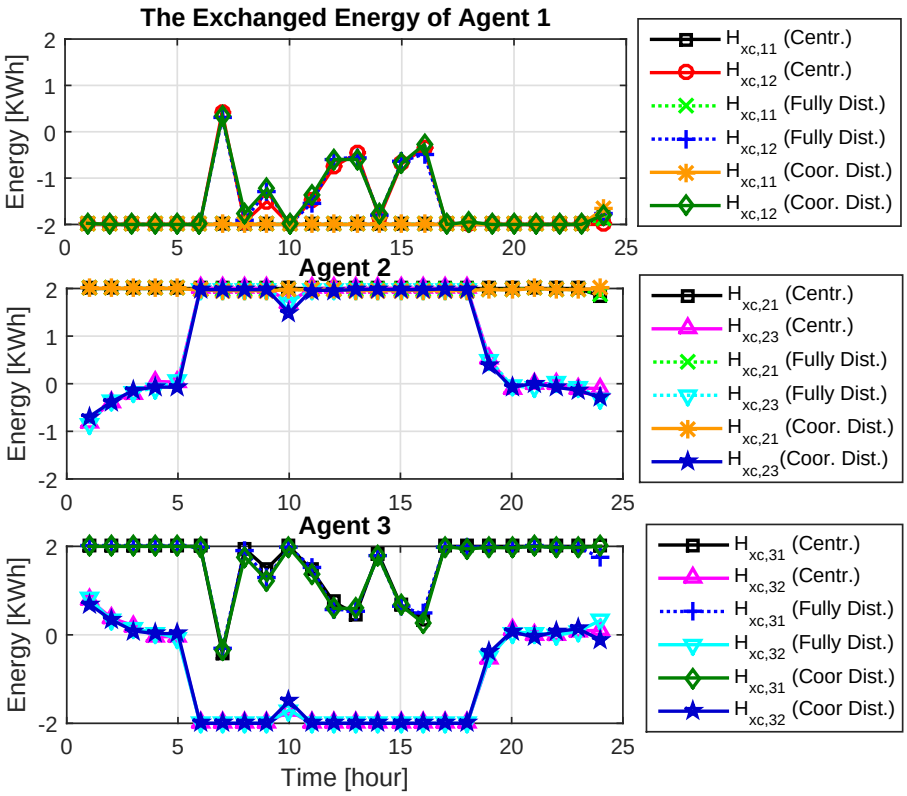


Figure 2-7: The thermal energy exchanged between agents.

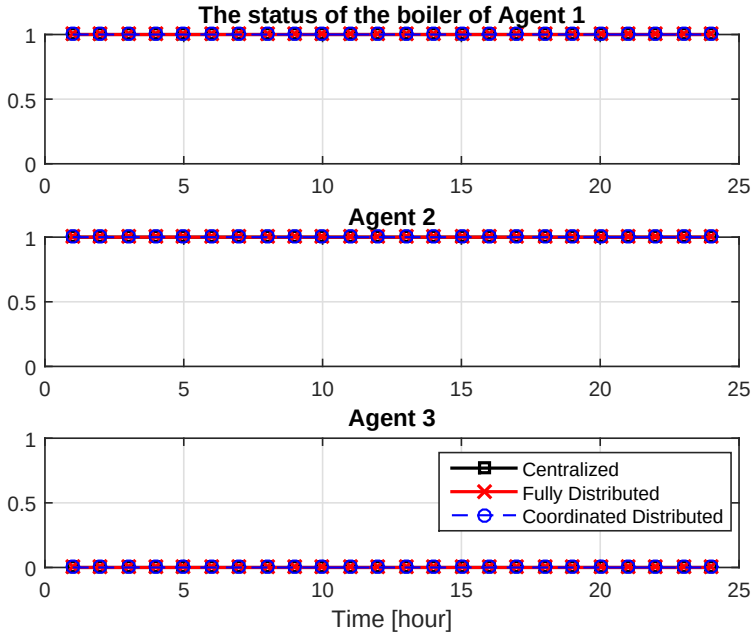


Figure 2-8: The status of the boilers.

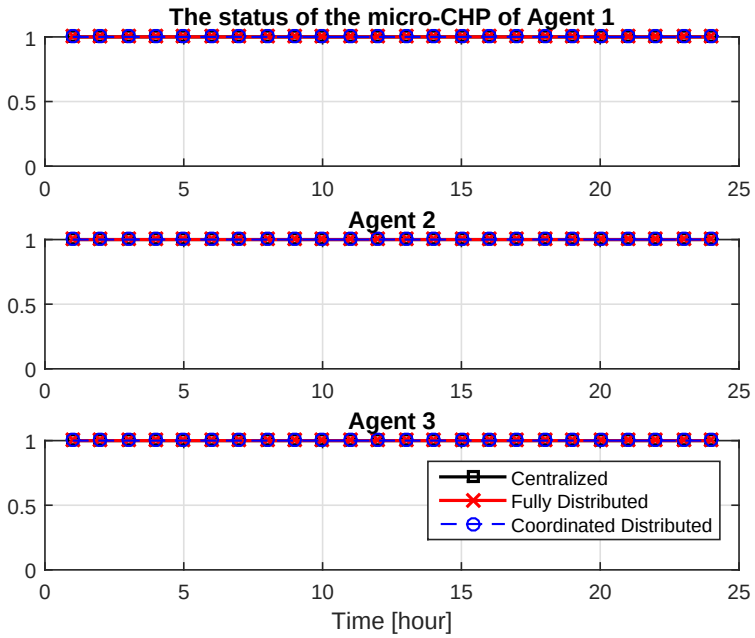


Figure 2-9: The status of the micro-CHPs.

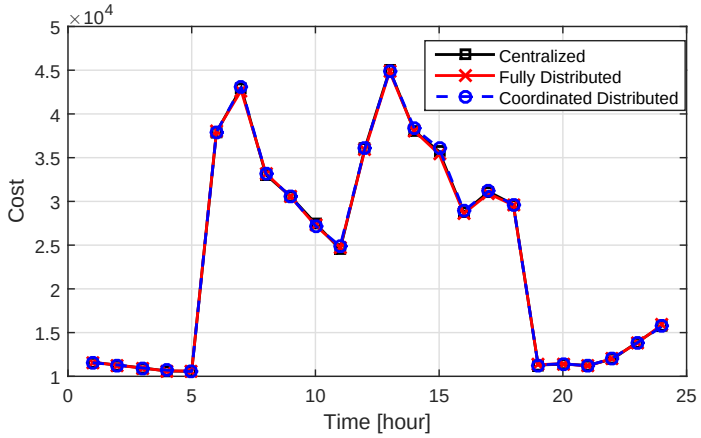


Figure 2-10: The costs at each sampling time.

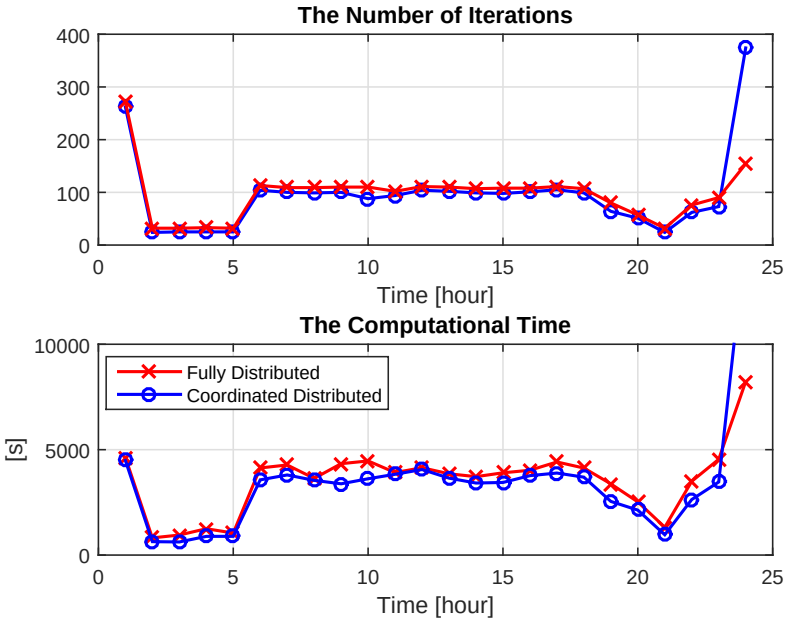


Figure 2-11: The number of iterations and the computational time required by the distributed approaches at each sampling time.

Chapter 3

Coordination in Thermal Grids with Seasonal Storages

In the second case study, we aim to exploit the advantage of having seasonal storages in thermal grids. The grids that are considered consist of N number of agents (buildings and houses), each of which has local thermal energy demand that can be fulfilled by its local production unit, its seasonal storage and/or an external party. In these grids, it is assumed that the production unit contains not only a boiler which supplies thermal energy but also a chiller which absorbs it when needed, for instance during the summer. Moreover, the seasonal storage can also perform both functions and become a cheaper alternative. However, such storage systems have dynamics and constraints which have to be taken into account. In addition, we do not consider micro-CHP, electrical power demand and the possibility of exchanging energy between agents since the emphasis of this chapter is on exploring the benefits of seasonal storages.

This chapter provides a brief description of yearly-based thermal energy demand profile in Section 3-1. Then, Section 3-2 describes the grid model including seasonal storages whereas Section 3-3 discusses the problem formulation which we wish to solve. In Section 3-4, we discuss the Hierarchical MPC which is proposed to solve the problem while in Section 3-5 we provide the numerical study which shows the performance of our proposed method.

3-1 Annual Thermal Energy Demand Profile

Thermal energy demand fluctuates over a year and in fact could be negative. Negative thermal demand means that the agent has some amount of energy that needs to be absorbed so that the thermal energy balance of the agent is satisfied. It can also be perceived as cooling demand while the positive values are viewed as heating demand. Figure 3-1 is given in order to illustrate the thermal energy demand profile of an agent in two years which is obtained from the Low Energy Architecture (LEA) model and real weather data in the Netherlands from mid 2010 until mid 2012. We can infer from the first plot of the figure that

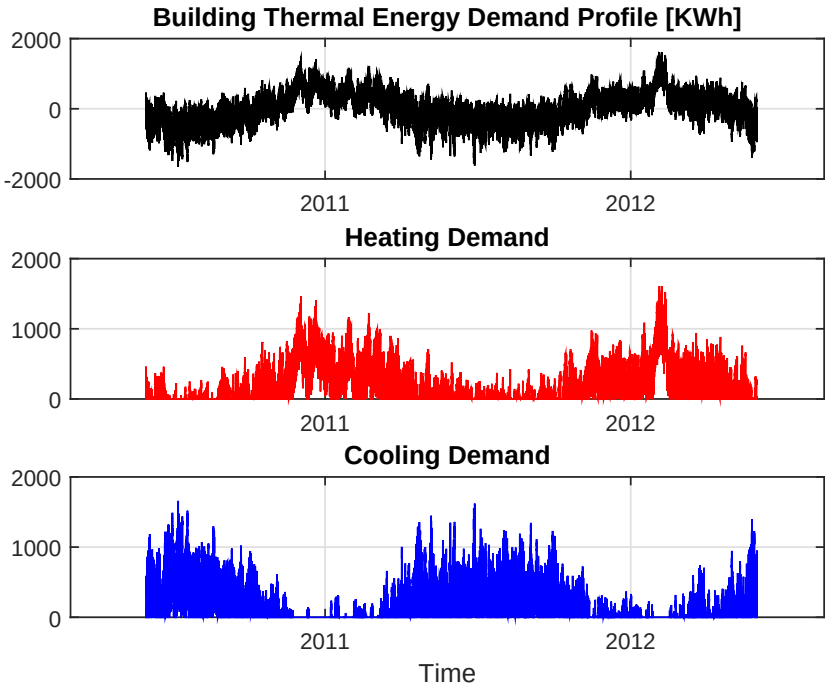


Figure 3-1: Thermal energy demand profile of a building. The first plot show the total demand over 5 years while the last two plots separately show the heating and cooling demands.

the fluctuation of the total demand has a yearly pattern. It is also seen that the agent has cooling demand mostly during the summers due to the fact that in this period the outdoor air temperature was relatively higher than the indoor air temperature which is set to be 19^oC and the solar radiation was also higher than the other months. On the other hand, heating demand dominates in the winter for the opposite reasons.

3-2 Modeling the Thermal Grids with Seasonal Storages

The model of the thermal grids, which, in this chapter, are considered to have seasonal storages, is built based on the dynamics of the seasonal storages and the imbalance errors as well as the operational constraints of the components. In this section, the focus is on describing the dynamics and the constraints of the seasonal storages as well as the dynamics of the imbalance errors which have to be slightly adjusted from the previous case study.

3-2-1 Dynamics and Constraints of Seasonal Storages

In general, the purpose of seasonal storages is to store thermal energy during the summer in order to heat buildings in the colder months [40]. They could increase the efficiency of the thermal grid in term of the usage of the production units since we can see the storages as additional energy sources which supports the system. As stated in [40], the ground can be used as a seasonal storage with which there are two methods of exchanging thermal energy:

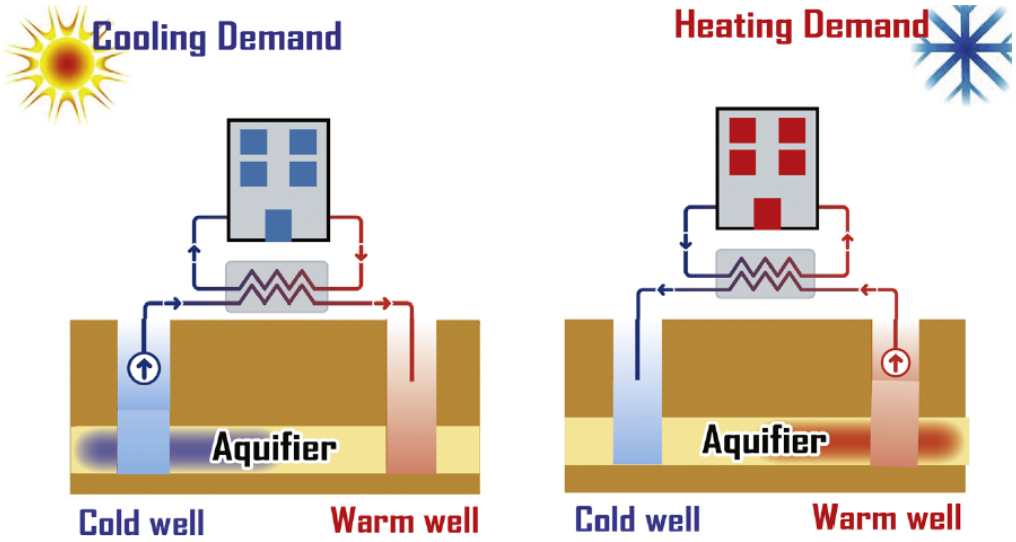


Figure 3-2: Operational modes of ATES. Figure is taken from [41].

through advection in aquifers and conduction using boreholes that are known as Aquifer Thermal Energy Storage (ATES) and Borehole Thermal Energy Storage (BTES), respectively. In our study, we consider the ATES systems which is largely used in the Netherlands. Such storage consists of warm and cold wells in which warm and cold water are stored. The typical temperature of the water in the warm wells and the cold wells are 14-17 ^{o}C and 6-9 ^{o}C, respectively, while the aquifer ambient water temperature is 10-13 ^{o}C [40]. Thus, the amount of supply that is stored in the wells can be considered to be proportional with the temperature difference and the volume of the water.

There are two operating conditions of the ATES systems which is illustrated in Figure 3-2. When the building that is connected with the ATES system requires heating supply, for instance during a cold season, the storage transfers the warm water from the warm wells to the cold wells by passing through a heat exchanger so that the building absorbs the thermal energy contained in the water. This causes the temperature of the water that will be stored in the cold wells drops. In this situation, it can be said that the water volume in the warm wells decreases while the water volume in the cold wells increases. On the other hand, during a warm season, where the building has cooling demand, the water in the cold well is extracted to absorb the thermal energy that is contained in the building. This results in the increasing of the water temperature which is then stored in the warm well. Moreover, in an ATES system heat pumps are also necessary to boost the temperature of the water from the warm well since it is not high enough to heat the building [40].

The first attempt to model an ATES system for an optimal control purpose, to the best of our knowledge, has been addressed in [42]. Here, we extend the the proposed model in [42] by using the similar idea in [43]. In order to describe the operation of the seasonal storage of agent i mathematically, we consider the water volume of the wells ($V_{\mathbf{h},i}(k), V_{\mathbf{c},i}(k) \in \mathbb{R}_+$) and the level of heating and cooling supplies ($s_{\mathbf{h},i}(k), s_{\mathbf{c},i}(k) \in \mathbb{R}_+$) as the states which have

the following discrete-time dynamic:

$$V_{\mathbf{h},i}(k+1) = V_{\mathbf{h},i}(k) - (u_{\mathbf{sh},i}(k) - u_{\mathbf{sc},i}(k)), \tag{3-1}$$
$$V_{\mathbf{c},i}(k+1) = V_{\mathbf{c},i}(k) + (u_{\mathbf{sh},i}(k) - u_{\mathbf{sc},i}(k)), \tag{3-2}$$
$$s_{\mathbf{h},i}(k+1) = s_{\mathbf{h},i}(k) - \alpha_{\mathbf{h},i}(u_{\mathbf{sh},i}(k) - u_{\mathbf{sc},i}(k)), \tag{3-3}$$
$$s_{\mathbf{c},i}(k+1) = s_{\mathbf{c},i}(k) + \alpha_{\mathbf{c},i}(u_{\mathbf{sh},i}(k) - u_{\mathbf{sc},i}(k)), \tag{3-4}$$

where the subscripts **h** and **c** refers to the warm and cold well respectively, $u_{\mathbf{sh},i}(k), u_{\mathbf{sc},i}(k) \in \mathbb{R}_+$ denote the volume of the water that is transferred for heating and cooling purposes while $\alpha_{\mathbf{h},i}$ and $\alpha_{\mathbf{c},i}$ are the volume-to-energy coefficient. Note that $u_{\mathbf{sh},i}(k)$ and $u_{\mathbf{sc},i}(k)$ are the decision variables of the controller which can be translated as the operating point of the pumps that are used to extract water from the well. Additionally, the volume-to-energy coefficients are computed as the following equations:

$$\alpha_{\mathbf{h},i} = \rho_{\mathrm{w}} c_{\mathrm{w}} \Delta T_{\mathbf{h},i}, \tag{3-5}$$
$$\alpha_{\mathbf{c},i} = \rho_{\mathrm{w}} c_{\mathrm{w}} \Delta T_{\mathbf{c},i}, \tag{3-6}$$

where $\rho_{\mathrm{w}}[\mathrm{kg/m^3}]$ is the water density, $c_{\mathrm{p,w}}$ $[\mathrm{Jkg^{-1}K^{-1}}]$ is the specific heat capacity of water, $\Delta T_{\mathbf{h},i}[\mathrm{K}]$ is the water temperature difference of the warm well and the aquifer ambient temperature and $\Delta T_{\mathbf{c},i}[\mathrm{K}]$ is the water temperature difference of the cold well and the aquifer ambient temperature.

Water in the warm and the cold wells may be mixed since it is stored in aquifers [40]. This causes the reducing of heating and cooling supplies. In current practice, the distance of the wells are designed to be three times than the thermal radius of the well to prevent the breakthrough [40]. Based on the mathematical formulation of the dynamic of one agent, we can infer that if one well, for instance the cold well grows, then the other well will reduce in the same rate. Thus we can say that the volume of the hot and cold well of one agent will not overlap each other. However, the overlap may happen when there exist neighboring agents which have a seasonal storage that has different growing/reducing rate. Therefore, we propose to impose additional constraints such that the controller can intelligently avoid this situation. They are represented in Equation (3-7) - (3-8), in which $\bar{V}_{\mathbf{hc},ij}$ is the maximum limit of the total volume of the warm well of agent i and the neighboring cold well which belongs to agent j and $\bar{V}_{\mathbf{ch},ig}$ is the maximum limit of the total volume of the cold well of agent i and the neighboring warm well which belongs to neighbor g.

$$V_{\mathbf{h},i}(k) + V_{\mathbf{c},j}(k) \leq \bar{V}_{\mathbf{hc},ij}, \quad \forall j \in \mathcal{N}_{-i,\mathbf{c}}, \tag{3-7}$$
$$V_{\mathbf{c},i}(k) + V_{\mathbf{h},g}(k) \leq \bar{V}_{\mathbf{ch},ig}, \quad \forall g \in \mathcal{N}_{-i,\mathbf{h}}, \quad \forall i \in \mathcal{N}, \tag{3-8}$$

where $\mathcal{N}_{-i,\mathbf{c}} \subseteq \mathcal{N} \backslash \{-i\}$ is the set of the cold wells which belong to the neighbors of agent i while $\mathcal{N}_{-i,\mathbf{h}} \subseteq \mathcal{N} \backslash \{-i\}$ is the set of the warm wells which belong to the neighbors of agent i. Furthermore, we define that $\mathcal{N}_{-i,\mathbf{h}} \cup \mathcal{N}_{-i,\mathbf{c}} = \mathcal{N}_{-i}$.

Another constraint that is addressed in order to ensure the sustainability of ATES systems is as follows. The energy that is injected in and extracted from them over some period of time must be balance [41]. This can be translated as the following equality constraint:

$$s_{\mathbf{h},i}(k) + s_{\mathbf{c},i}(k) = \bar{s}_{\mathbf{hc},i}, \tag{3-9}$$

where $\bar{s}_i$ refers to the maximum amount of supply in one well when the other well is empty. However, the conditions between wells varies over time and this could lead to the violation of this constraint. Therefore we introduce $e_i(k)$ as an auxiliary decision variable which allows some level of violation of the constraint and relax it into

$$s_{\mathbf{h},i}(k) + s_{\mathbf{c},i}(k) \leq \bar{s}_{\mathbf{hc},i} + e_i(k). \tag{3-10}$$

3-2-2 Dynamics of the Imbalance Errors

In this case study, heating and cooling demands are distinguished and denoted as $h_{\mathbf{d}}$ and $c_{\mathbf{d}}$ which are non-negative, *i.e.* $h_{\mathbf{d}}, c_{\mathbf{d}} \geq 0$. Since now we have separated heating and cooling demands, we also need to define two different imbalance errors which correspond to them. In essence, the imbalance error indicates the difference between the supply and the demand of energy, therefore we define the imbalance errors as follows:

$$x_{\mathbf{h},i}(k) = h_{\mathbf{buf},i}(k) - h_{\mathbf{d},i}(k), \tag{3-11}$$
$$x_{\mathbf{c},i}(k) = c_{\mathbf{buf},i}(k) - c_{\mathbf{d},i}(k), \tag{3-12}$$

where $x_{\mathbf{h},i}(k)$ and $x_{\mathbf{c},i}(k)$ are the imbalance errors of agent i at sampling time k that are related to heating and cooling, respectively, $h_{\mathbf{buf},i}(k)$ and $c_{\mathbf{buf},i}(k)$ denote the level of heating and cooling supplies that are currently available in the buffers. Furthermore, the dynamics of the buffers are expressed as

$$h_{\mathbf{buf},i}(k+1) = \eta_{\mathbf{buf},i}\left(x_{\mathbf{h},i}(k) + h_{\mathbf{b},i}(k) + h_{\mathbf{im},i}(k) + \alpha_{\mathbf{cop}} h_{\mathbf{a},i}(k)\right), \tag{3-13}$$
$$c_{\mathbf{buf},i}(k+1) = \eta_{\mathbf{buf},i}\left(x_{\mathbf{c},i}(k) + c_{\mathbf{ch},i}(k) + c_{\mathbf{im},i}(k) + c_{\mathbf{a},i}(k)\right). \tag{3-14}$$

Equation (3-13) represents the dynamics of heating supply in the buffer where $\eta_{\mathbf{buf},i} \in [0,1]$ is the efficiency of the buffer, $\alpha_{\mathbf{hp}}$ is the coefficient which is related to the performance of heat pump, $h_{\mathbf{b},i}(k) \in \mathbb{R}_+$ is the heat supplied by the local boiler, $h_{\mathbf{im},i}(k) \in \mathbb{R}_+$ is the imported heat from an external party and $h_{\mathbf{a},i}(k) \in \mathbb{R}_+$ is the heating supply that is taken from the seasonal storage. Whereas, Equation (3-14) provides the dynamics of the cooling supply in the buffer in which $c_{\mathbf{ch},i}(k), c_{\mathbf{im},i}(k), c_{\mathbf{a},i}(k) \in \mathbb{R}_+$ indicate the cooling supplies that are taken from the chiller, imported from an external party and given by the seasonal storage, respectively. The heating and cooling supplies from the storage, $h_{\mathbf{a},i}(k)$ and $c_{\mathbf{a},i}(k)$, are proportional to the water volume that is transferred between the wells of the storage, *i.e.*:

$$h_{\mathbf{a},i}(k) = \alpha_{\mathbf{h},i} u_{\mathbf{sh},i}(k), \tag{3-15}$$
$$c_{\mathbf{a},i}(k) = \alpha_{\mathbf{c},i} u_{\mathbf{sc},i}(k), \tag{3-16}$$

3-3 Problem Formulation

In this part, we formulate the optimization problem of the production planning for the grids with seasonal storages. Based on Equation (3-11) - (3-14), the dynamics of the imbalance error of the system can be represented in as a state-space description as follows

$$\boldsymbol{x}_i(k+1) = A_i \boldsymbol{x}_i(k) + B_i \boldsymbol{u}_i(k) + H_i \boldsymbol{w}_i(k), \tag{3-17}$$

where

$$\boldsymbol{x}_i(k) = \begin{bmatrix} x_{\mathbf{h},i}(k) & x_{\mathbf{c},i}(k) \end{bmatrix}^T \in \mathbb{R}^2,$$
$$\boldsymbol{u}_i(k) = \begin{bmatrix} \boldsymbol{u}_{h,i}^T(k) & \boldsymbol{u}_{c,i}^T(k) & \boldsymbol{u}_{s,i}^T(k) & \boldsymbol{u}_{a,i}^T(k) \end{bmatrix}^T \in \mathbb{R}^9,$$
$$\boldsymbol{u}_{h,i}(k) = \begin{bmatrix} h_{\mathbf{b},i}(k) & h_{\mathbf{im},i}(k) \end{bmatrix}^T \in \mathbb{R}^2,$$
$$\boldsymbol{u}_{c,i}(k) = \begin{bmatrix} c_{\mathbf{ch},i}(k) & c_{\mathbf{im},i}(k) \end{bmatrix}^T \in \mathbb{R}^2,$$
$$\boldsymbol{u}_{s,i}(k) = \begin{bmatrix} u_{\mathbf{sh},i}(k) & u_{\mathbf{sc},i}(k) \end{bmatrix}^T \in \mathbb{R}^2,$$
$$\boldsymbol{u}_{a,i}(k) = \begin{bmatrix} u_{\mathbf{b},i}^{\mathbf{su}}(k) & u_{\mathbf{ch},i}^{\mathbf{su}}(k) & e_i(k) \end{bmatrix}^T \in \mathbb{R}^3,$$
$$\boldsymbol{w}_i = \begin{bmatrix} h_{\mathbf{d},i}(k+1) & c_{\mathbf{d},i}(k+1) \end{bmatrix}^T \in \mathbb{R}^2,$$
$$A_i = \eta_{\mathbf{buf},i} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad B_i = \eta_{\mathbf{buf},i} \begin{bmatrix} \mathbf{1}^{1\times 2} & \mathbf{0}^{1\times 2} & \alpha_{\mathbf{cop}} & 0 & \mathbf{0}^{1\times 3} \\ \mathbf{0}^{1\times 2} & \mathbf{1}^{1\times 2} & 0 & 1 & \mathbf{0}^{1\times 3} \end{bmatrix}, \quad H_i = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Meanwhile, based on Equation (3-1) - (3-4) and (3-15)-(3-16), the dynamics of the seasonal storages can be simply described as

$$\boldsymbol{x}_{\mathbf{a},i}(k+1) = \Phi_i \boldsymbol{x}_{\mathbf{a},i}(k) + \Gamma_i \boldsymbol{u}_i(k), \tag{3-18}$$

where the state of the seasonal storages is defined as

$$\boldsymbol{x}_{\mathbf{a},i}(k) = \begin{bmatrix} V_{\mathbf{h},i}(k) & V_{\mathbf{c},i}(k) & s_{\mathbf{h},i}(k) & s_{\mathbf{c},i}(k) \end{bmatrix}^T \in \mathbb{R}^4,$$

and the state-space matrices Φ_i and Γ are defined as:

$$\Phi_i = I^{4\times 4}, \quad \Gamma_i = \begin{bmatrix} \mathbf{0}^{1\times 4} & -1 & 1 & \mathbf{0}^{1\times 3} \\ \mathbf{0}^{1\times 4} & 1 & -1 & \mathbf{0}^{1\times 3} \\ \mathbf{0}^{1\times 4} & -\alpha_{\mathbf{ah},i} & \alpha_{\mathbf{ah},i} & \mathbf{0}^{1\times 3} \\ \mathbf{0}^{1\times 4} & \alpha_{\mathbf{ac},i} & -\alpha_{\mathbf{ac},i} & \mathbf{0}^{1\times 3} \end{bmatrix}.$$

In this state-space description, the input of the state space equation or the decision vector of variables are denoted by $\boldsymbol{u}_i$ which consists $\boldsymbol{u}_{h,i}(k)$, $\boldsymbol{u}_{c,i}(k)$, $\boldsymbol{u}_{s,i}(k)$ and $\boldsymbol{u}_{a,i}(k)$. The vector $\boldsymbol{u}_{h,i}(k)$ and $\boldsymbol{u}_{c,i}(k)$ contain the manipulated variables that are related to the heating and cooling supplies which come from the boilers, the chillers, and the external party, $\boldsymbol{u}_{s,i}(k)$ consists of the manipulated variable related to the seasonal storages while $\boldsymbol{u}_{a,i}(k)$ consists of the auxiliary variable necessary for the start up cost constraints and the constraints of the seasonal storages. We also introduce $\boldsymbol{v}_i(k) = \begin{bmatrix} v_{\mathbf{b},i}(k) & v_{\mathbf{ch},i}(k) \end{bmatrix}^T \in \{0,1\} \times \{0,1\}$ as the vector of the binary decision variables which represents the status of the production units.

In this case, we again deal with the production planning problem in which we want to set the optimal operating condition of the production unit of each agent at every hour. We consider that each agent has the startup cost and the capacity constraints of the production of the supply. Moreover, the reliability is assured by imposing non-negativity constraints on the

imbalance errors and their predictions. These become chance constraints due to the existence of the uncertain demand. The cost function that is minimized is a quadratic cost with respect to the imbalance error and the decision variables as

$$J_i(k) = \boldsymbol{x}_i^T(k) Q_i \boldsymbol{x}_i(k) + \boldsymbol{u}_i^T(k) R_i \boldsymbol{u}_i(k), \quad \forall i \in \mathcal{N}, \tag{3-19}$$

in which $Q_i, R_i > 0$, where

$$R_i = \text{diag}([r_{\mathbf{b},i} \;\; r_{\mathbf{im},i} \;\; r_{\mathbf{ch},i} \;\; r_{\mathbf{im},i} \;\; r_{\mathbf{ush},i} \;\; r_{\mathbf{usc},i} \;\; 1 \;\; 1 \;\; 1]),$$

where $r_{\mathbf{b},i}, r_{\mathbf{ch},i}, r_{\mathbf{im},i}, r_{\mathbf{ush},i}$ and $r_{\mathbf{usc},i}$ denote the weight of $h_{\mathbf{b},i}, c_{\mathbf{ch},i}, h_{\mathbf{im},i}/c_{\mathbf{im},i}, u_{\mathbf{sh},i}$ and $u_{\mathbf{sh},i}$ respectively. By having the prediction horizon of N_h, we have the following optimization problem that need to be solved at every sampling time k:

$$\underset{\{\boldsymbol{u}_i(k:k+N_h-1), \boldsymbol{v}_i(k:k+N_h-1)\}_{i=1}^N}{\text{minimize}} \quad \sum_{t=k}^{k+N_h-1} \sum_{i=1}^{N} J_i(t) \tag{3-20}$$

subject to:

1. The dynamics of the imbalance error:

$$\boldsymbol{x}_i(t+1) = A_i \boldsymbol{x}_i(t) + B_i \boldsymbol{u}_i(t) + H_i \boldsymbol{w}_i(t), \quad \forall i \in \mathcal{N}, \quad \forall t \in \{k, \cdots, N_h + k - 1\}. \tag{3-21}$$

2. The dynamics of the seasonal storage:

$$\boldsymbol{x}_{\mathbf{a},i}(k+1) = \Phi_i \boldsymbol{x}_{\mathbf{a},i}(k) + \Gamma_i \boldsymbol{u}_i(k), \; \forall i \in \mathcal{N}, \; \forall t \in \{k, \cdots, N_h + k - 1\}. \tag{3-22}$$

3. The startup cost constraints:

$$c_i(t)^{\mathbf{su}} \geq \Lambda_i^{\mathbf{su}} (v_i(t) - v_i(t-1)) \;, \quad c_{i,t}^{\text{su}} \geq 0, \quad \forall i \in \mathcal{N}, \quad \forall t \in \{k, \cdots, N_h + k - 1\}, \tag{3-23}$$

   where Λ_i^{su} is a diagonal matrix in which the diagonal elements are the startup costs of the boiler and the chiller owned by an agent.

4. The capacity constraints:

$$v_{\mathbf{b},i}(t) h_{\mathbf{b},i}^{\min} \leq h_{\mathbf{b},i}(t) \leq v_{\mathbf{b},i}(t) h_{\mathbf{b},i}^{\max} \;, \tag{3-24}$$
$$h_{\mathbf{im},i}^{\min} \leq h_{\mathbf{im},i}(t) \leq h_{\mathbf{im},i}^{\max} \;, \tag{3-25}$$
$$v_{\mathbf{ch},i}(t) c_{\mathbf{ch},i}^{\min} \leq c_{\mathbf{ch},i}(t) \leq v_{\mathbf{ch},i}(t) c_{\mathbf{ch},i}^{\max} \;, \tag{3-26}$$
$$c_{\mathbf{im},i}^{\min} \leq c_{\mathbf{im},i}(t) \leq c_{\mathbf{im},i}^{\max} \;, \tag{3-27}$$
$$u_{\mathbf{sh},i}^{\min} \leq u_{\mathbf{sh},i}(t) \leq u_{\mathbf{sh},i}^{\max} \;, \tag{3-28}$$
$$u_{\mathbf{sc},i}^{\min} \leq u_{\mathbf{sc},i}(t) \leq u_{\mathbf{sc},i}^{\max} \;, \quad \forall i \in \mathcal{N}, \quad \forall t \in \{k, \cdots, N_h + k - 1\}. \tag{3-29}$$

5. The local constraints of the seasonal storages:

$$s_{\mathbf{h},i}(t+1) + s_{\mathbf{c},i}(t+1) \leq \bar{s}_{\mathbf{hc},i} + e_i(k) \tag{3-30}$$
$$s_{\mathbf{h},i}(t+1) \geq 0, \quad s_{\mathbf{c},i}(t+1) \geq 0 \quad \forall i \in \mathcal{N}, \quad \forall t \in \{k, \cdots, N_h + k - 1\}. \tag{3-31}$$

6. The coupling constraints of the seasonal storages:

$$V_{\mathbf{h},i}(t+1) + V_{\mathbf{c},j}(t+1) \leq \bar{V}_{\mathbf{hc},ij}, \quad \forall j \in \mathcal{N}_{-i,\mathbf{c}} \tag{3-32}$$

$$V_{\mathbf{c},i}(t+1) + V_{\mathbf{h},g}(t+1) \leq \bar{V}_{\mathbf{ch},ig}, \quad \forall g \in \mathcal{N}_{-i,\mathbf{h}}, \quad \forall i \in \mathcal{N}, \quad \forall t \in \{k, \cdots, N_h + k - 1\}. \tag{3-33}$$

7. The probabilistic constraints

$$\mathbb{P}(\boldsymbol{x}(t+1) \geq 0, \quad \forall i \in \mathcal{N}, \quad \forall t \in \{k, \cdots, N_h + k - 1\}) \geq 1 - \varepsilon, \tag{3-34}$$

where ε is the level of violation.

This optimization problem can be written compactly as follows

$$\underset{\{\tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{v}}_i\}_{i=1}^N}{\text{minimize}} \quad \sum_{i=1}^{N} \boldsymbol{J}_i(\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{u}}_i) \tag{3-35a}$$

$$\text{subject to} \quad \mathbb{P}(\mathbf{A}\boldsymbol{x}_i(k) + \mathbf{B}_i\tilde{\boldsymbol{u}}_i + \mathbf{H}_i\tilde{\boldsymbol{w}}_i \geq 0, \forall i \in \mathcal{N}) \geq 1 - \varepsilon, \tag{3-35b}$$

$$\mathbf{E}_i\tilde{\boldsymbol{u}}_i + \mathbf{F}_i\tilde{\boldsymbol{v}}_i + \mathbf{P}_i \leq 0, \tag{3-35c}$$

$$\boldsymbol{\Phi}_i\boldsymbol{x}_{\mathbf{a},i}(k) + \boldsymbol{\Gamma}_i\tilde{\boldsymbol{u}}_i \geq 0, \tag{3-35d}$$

$$\boldsymbol{\Psi}\left(\boldsymbol{\Phi}_i\boldsymbol{x}_{\mathbf{a},i}(k) + \boldsymbol{\Gamma}_i\tilde{\boldsymbol{u}}_i\right) - \boldsymbol{\Psi}_e\tilde{\boldsymbol{u}}_i \leq \bar{\boldsymbol{s}}_{\mathbf{hc},i}, \tag{3-35e}$$

$$\mathbf{G}_{ii}\left(\boldsymbol{\Phi}_i\boldsymbol{x}_{\mathbf{a},i}(k) + \boldsymbol{\Gamma}_i\tilde{\boldsymbol{u}}_i\right) + \sum_{j \in \mathcal{N}_{-i}} \mathbf{G}_{ij}\left(\boldsymbol{\Phi}_j\boldsymbol{x}_{\mathbf{a},j}(k) + \boldsymbol{\Gamma}_{s,j}\tilde{\boldsymbol{u}}_j\right) \leq \boldsymbol{\Theta}_i, \quad \forall i \in \mathcal{N}, \tag{3-35f}$$

in which

$$\boldsymbol{J}_i(\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{u}}_i) = \tilde{\boldsymbol{x}}_i^T\boldsymbol{Q}_i\tilde{\boldsymbol{x}}_i + \tilde{\boldsymbol{u}}_i^T\boldsymbol{R}_i\tilde{\boldsymbol{u}}_i, \tag{3-36}$$

by denoting that $\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{w}}_i, \tilde{\boldsymbol{v}}_i, \mathbf{A}_i, \mathbf{B}_i, \mathbf{H}_i, \boldsymbol{\Psi}_i$ and $\boldsymbol{\Gamma}_i$ are defined in Appendix B-1 and

$$\Psi = \begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix}, \quad \Psi_e = \begin{bmatrix} \mathbf{0}^{1\times 8} & 1 \end{bmatrix}.$$

Moreover $\boldsymbol{\Psi} \in \mathbb{R}^{N_h \times 4N_h}$ and $\boldsymbol{\Psi}_e \in \mathbb{R}^{N_h \times N_u N_h}$ are block diagonal matrices of Ψ and Ψ_e, while $\mathbf{G}_{ii} \in \mathbb{R}^{|\mathcal{N}_{-i}|N_h \times 4N_h} \ \forall i \in \mathcal{N}$ are block diagonal matrices of $G_{ii} \in \mathbb{R}^{|\mathcal{N}_{-i}| \times 4}$ and $\mathbf{G}_{ij} \in \mathbb{R}^{|\mathcal{N}_{-i}|N_h \times 4N_h} \ \forall j \in \mathcal{N}_{-i} \ \forall i \in \mathcal{N}$ are block diagonal matrices of $G_{ij} \in \mathbb{R}^{|\mathcal{N}_{-i}| \times 4}$. G_{ii} and G_{ij} are constructed such that the coupling constraints, *i.e.* Equation (3-32) - (3-33), are satisfied (see an example of the construction of these matrices for the numerical study in Appendix B-5). Furthermore, $\bar{\boldsymbol{s}}_{\mathbf{hc},i} \in \mathbb{R}^{N_h}$ is a vector whose elements are $\bar{s}_{\mathbf{hc},i}$, $\boldsymbol{\Theta}_i \in \mathbb{R}^{N_h|\mathcal{N}_{-i}|}$ consists of $\Theta_i \in \mathbb{R}^{|\mathcal{N}_{-i}|}$, elements of which are $\bar{V}_{\mathbf{hc},ij}$ and $\bar{V}_{\mathbf{ch},ig}$.

Problem (3-35) is a mixed-integer chance-constrained program in which Equation (3-35b) denotes the chance constraint and Equation (3-35c) collects the local constraints (3-23) - (3-29) such that $\mathbf{E}_i$, $\mathbf{F}_i$,$\mathbf{P}_i$ have appropriate dimensions and correspond to them as explained in Appendix B-4. Moreover, Equation (3-35d) is the non-negativity constraint of the predictions of the supply level of the hot and cold storage and comes from (3-31), Equation (3-35e) is derived from the join maximum capacity limit of the hot and cold storage of the same agent (3-30) and Equation (3-35f) is the coupling constraint between agents which comes from Equation (3-32) and (3-33).

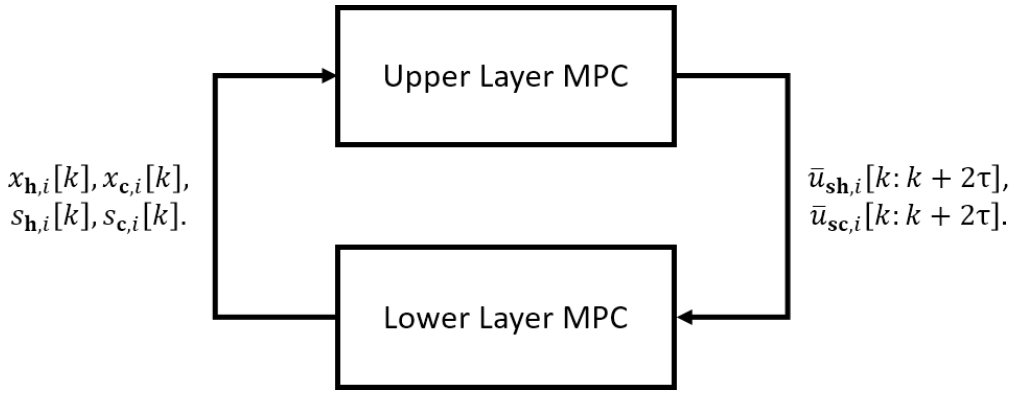


Figure 3-3: The proposed hierarchical MPC scheme.

3-4 Hierarchical MPC

The availability of the supply of the seasonal storages may not be able to fulfill the demand over the year. Thus the controller that is proposed should be able to capture the yearly-based thermal demand profile in order to optimize the usage of the supply of the seasonal storages. An obvious way to achieve it is by having an appropriate prediction horizon. However, since the controller sets the hourly operating point of the production units, the problem becomes intractable if we impose a very long prediction horizon, for instance, one year, *i.e.* $N_h = 8760$, by assuming that one year is 365 days. The main reason which causes the intractability is simply the number of decision variables which is too large. Moreover, if we consider the robust randomized Model Predictive Control (MPC) approach as the method to solve this problem, not only the part of solving the robust optimization but the computation of the bounds will become intractable.

Therefore, we propose a hierarchical MPC scheme to tackle this issue. It consists of two layer of MPC controllers as depicted in Figure 3-3. The task of the upper layer is to provide the maximum amount of supply that can be taken from the seasonal storage at each hour while the second layer computes the hourly set points. The aim of providing the upper bounds of $u_{\mathbf{sh},i}$ and $u_{\mathbf{sc},i}$ from the upper layer is to ensure that the supplies of the seasonal storages are not greedily used, which could happen, since it can be seen as free sources. When computing the upper bounds, the upper layer captures the yearly-based demand profile of the grid by considering a larger sampling time, which here we define as one month, and monthly demands. In addition, the upper layer provides the hourly upper bounds for one month and does the optimization only in less frequent manner than the lower layer which in this case is at the beginning of the month.

3-4-1 Formulation of the Upper and the Lower Layer MPC

In the upper layer, we consider the production planning problem in a larger time scale, *i.e.,* the sampling time is one month. However, since the goal of this part is to provide the upper bound of $u_{\mathbf{sh},i}$ and $u_{\mathbf{sc},i}$, we define the control variable $\boldsymbol{u}_{\mathbf{u},i}(m)$, in which the subscript $\mathbf{u}$ implies that it belongs to the upper layer problem, as

$$\boldsymbol{u}_{\mathbf{u},i}(m) = \begin{bmatrix} u_{\mathbf{oh,u},i}(m) & u_{\mathbf{oc,u},i}(m) & u_{\mathbf{sh,u},i}(m) & u_{\mathbf{sc,u},i}(m) & e_{\mathbf{u},i}(m) \end{bmatrix}^T \in \mathbb{R}^5, \tag{3-37}$$

where $u_{\mathbf{sh},\mathbf{u},i}(m)$ and $u_{\mathbf{sc},\mathbf{u},i}(m)$ denote the heating and the cooling supply of the seasonal storage of agent i whereas $u_{\mathbf{oh},\mathbf{u},i}(m)$ and $u_{\mathbf{oc},\mathbf{u},i}(m)$ refer to the heating and the cooling supply from other sources which can be the production units or imported from the external party. Additionally, $e_{\mathbf{u},i}(m)$ is the auxiliary variable which appears in the constraint of the ATES. Note that the time indexes used are noted as m since the sampling time of the upper layer MPC is different than the lower layer and it is larger, *i.e.*, $m = \tau k$ where $\tau > 1$. In our particular case, $\tau = 720$ since we assume one month is 720 hours.

The monthly optimization of the upper layer MPC with the prediction horizon of 12 months ($N_{h,\mathbf{u}} = 12$) is defined as the following:

$$\underset{\{\tilde{\boldsymbol{u}}_{\mathbf{u},i}\}_{i=1}^{N}}{\text{minimize}} \quad \sum_{i=1}^{N} \boldsymbol{J}_{\mathbf{u},i}(\tilde{\boldsymbol{u}}_{\mathbf{u},i}) \tag{3-38a}$$

$$\text{subject to} \quad \mathbb{P}(\mathbf{A}\boldsymbol{x}_{\mathbf{u},i}(m) + \mathbf{B}_{\mathbf{u},i}\tilde{\boldsymbol{u}}_{\mathbf{u},i} + \mathbf{H}_i\tilde{\boldsymbol{w}}_{\mathbf{u},i} \geq 0, \forall i \in \mathcal{N}) \geq 1 - \varepsilon, \tag{3-38b}$$

$$\boldsymbol{\Phi}_i\boldsymbol{x}_{\mathbf{a},\mathbf{u},i}(m) + \boldsymbol{\Gamma}_{\mathbf{u},i}\tilde{\boldsymbol{u}}_{\mathbf{u},i} \geq 0, \tag{3-38c}$$

$$\boldsymbol{\Psi}\left(\boldsymbol{\Phi}_i\boldsymbol{x}_{\mathbf{a},\mathbf{u},i}(m) + \boldsymbol{\Gamma}_{\mathbf{u},i}\tilde{\boldsymbol{u}}_{\mathbf{u},i}\right) - \boldsymbol{\Psi}_{e,\mathbf{u}}\tilde{\boldsymbol{u}}_{\mathbf{u},i} \leq \bar{\boldsymbol{s}}_{\mathbf{hc},i}, \tag{3-38d}$$

$$\mathbf{G}_{ii}\left(\boldsymbol{\Phi}_i\boldsymbol{x}_{\mathbf{a},\mathbf{u},i}(m) + \boldsymbol{\Gamma}_{\mathbf{u},i}\tilde{\boldsymbol{u}}_{\mathbf{u},i}\right) + \sum_{j\in\mathcal{N}_{-i}} \mathbf{G}_{ij}\left(\boldsymbol{\Phi}_j\boldsymbol{x}_{\mathbf{a},\mathbf{u},j}(m) + \boldsymbol{\Gamma}_{\mathbf{u},j}\tilde{\boldsymbol{u}}_{\mathbf{u},j}\right) \leq \boldsymbol{\Theta}_i, \tag{3-38e}$$

$$\tilde{\boldsymbol{u}}_{\mathbf{u},i} \geq 0, \quad \boldsymbol{x}_{\mathbf{u},i}(m) = \boldsymbol{x}_i(k), \quad \boldsymbol{x}_{\mathbf{a},\mathbf{u},i}(m) = \boldsymbol{x}_{\mathbf{a},\mathbf{u},i}(k), \quad \forall i \in \mathcal{N}. \tag{3-38f}$$

In this formulation, all the variables and matrices are in the compact form of the concatenated variables and matrices of the whole prediction horizon ($N_{h,\mathbf{u}}$) as explained in Appendix B. Moreover, $B_{\mathbf{u},i}$ which presents in $\mathbf{B}_{\mathbf{u},i}$ and $\Gamma_{\mathbf{u},i}$ which presents in $\boldsymbol{\Gamma}_{\mathbf{u},i}$ are defined as:

$$B_{\mathbf{u},i} = \begin{bmatrix} 1 & 0 & \alpha_{\mathbf{cop}} & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}, \quad \Gamma_{\mathbf{u},i} = \begin{bmatrix} \mathbf{0}^{1\times 2} & -1 & 1 & 0 \\ \mathbf{0}^{1\times 2} & 1 & -1 & 0 \\ \mathbf{0}^{1\times 2} & -\alpha_{\mathbf{ah},i} & \alpha_{\mathbf{ah},i} & 0 \\ \mathbf{0}^{1\times 2} & \alpha_{\mathbf{ac},i} & -\alpha_{\mathbf{ac},i} & 0 \end{bmatrix},$$

and $\boldsymbol{\Psi}_{e,\mathbf{u}} \in \mathbb{R}^{N_{h,\mathbf{u}}\times 5N_{h,\mathbf{u}}}$ is a block diagonal matrix of $\Psi_e = \begin{bmatrix}\mathbf{0}^{1\times 4} & 1\end{bmatrix}$.

The cost function (3-38a) is a sum of the quadratic cost of each agent which is defined as follows

$$J_{\mathbf{u},i} = \boldsymbol{x}_{\mathbf{u},i}^T(l)Q_{\mathbf{u},i}\boldsymbol{x}_{\mathbf{u},i}(l) + \boldsymbol{u}_{\mathbf{u},i}^T(l)R_{\mathbf{u},i}\boldsymbol{u}_{\mathbf{u},i}(l), \quad \forall i \in \mathcal{N}, \quad \forall l \in \{m, \cdots, m+N_{h,\mathbf{u}}-1\}, \tag{3-39}$$

where $Q_{\mathbf{u},i}$ and $R_{\mathbf{u},i}$ are positive definite matrices, *i.e.* $Q_{\mathbf{u},i}, R_{\mathbf{u},i} > 0$. Moreover, $R_{\mathbf{u},i}$ is a diagonal matrix where the weight of the supply of the seasonal storage is chosen to be smaller than the weight of the supply of the other sources. In this part, the weights become the tuning parameters which can be manipulated. Moreover, the constraints imposed in this layer are the chance constraint of the imbalance error (3-38b), the non-negativity constraint of the supply level of the storages (3-38c) and the joint maximum supply limit of the hot and cold storages of one agent (3-38d) as well as the joint maximum volume limit of the hot and cold storages of two agents (3-38e) which is the coupling constraint. Additionally, the upper layer requires the current states of the seasonal storages and the imbalance errors which are obtained from the lower layer and as stated in Equation (3-38f).

In the lower layer, the controller solves the original optimization problem (3-35) without the coupling constraint and with the additional constraint on the $u_{\mathbf{sh},i}(k)$ and $u_{\mathbf{sc},i}(k)$ as follows:

$$u_{\mathbf{sh},i}(t) \leq \bar{u}_{\mathbf{sh},i}(t), \quad \forall i \in \mathcal{N}, \quad \forall t \in \{k, \cdots, k+N_h-1\}, \tag{3-40}$$

$$u_{\mathbf{sc},i}(t) \leq \bar{u}_{\mathbf{sc},i}(t), \quad \forall i \in \mathcal{N}, \quad \forall t \in \{k, \cdots, k+N_h-1\}, \tag{3-41}$$

where $\bar{u}_{\mathbf{sh},i}$ and $\bar{u}_{\mathbf{sc},i}$ are obtained from the solution of the upper layer. Note that the coupling constraint is not taken into account in the lower layer because it has been included in the optimization in the upper layer. Thus, the upper bounds given by the upper layer are also expected to ensure the satisfaction of the coupling constraints. Therefore, the problem in the lower layer is separable and can be distributed to the agents. In the next part, we will discuss the method that we use to compute $\bar{u}_{\mathbf{sh},i}$ and $\bar{u}_{\mathbf{sc},i}$ at each hour from the solution of the upper layer MPC.

3-4-2 Obtaining the Maximum Limit of the Supply from the Seasonal Storages

Since the upper layer MPC solves the monthly production planning problem, its solutions, $u^{\star}_{\mathbf{sh,u},i}(m)$ and $u^{\star}_{\mathbf{sc,u},i}(m)$, are basically the total amount of heating and cooling supplies which can be taken from the seasonal storage i during the duration m. Now we want to allocate the amount of the available supplies to each sampling time k. Dividing $u^{\star}_{\mathbf{sh,u},i}(m)$ and $u^{\star}_{\mathbf{sc,u},i}(m)$ equally for each sampling time, *i.e.*

$$\bar{\boldsymbol{u}}_{\mathbf{sh},i}(k:k+\tau) = \frac{u^{\star}_{\mathbf{sh,u},i}(m)}{\tau}, \tag{3-42}$$

$$\bar{\boldsymbol{u}}_{\mathbf{sc},i}(k:k+\tau) = \frac{u^{\star}_{\mathbf{sc,u},i}(m)}{\tau}, \tag{3-43}$$

is not suitable since the demand at each sampling time varies. Therefore, in order to translate these values to become the upper bound as mentioned in (3-40) and (3-41), we use some knowledge of the demand profile to form two stochastic vectors. We obtain hourly $\bar{u}_{\mathbf{sh},i}$ and $\bar{u}_{\mathbf{sc},i}$ of one month as follows:

$$\bar{\boldsymbol{u}}_{\mathbf{sh},i}(k:k+\tau) = \frac{\boldsymbol{h}^{f}_{\mathbf{d},i}(k:k+\tau)}{\sum_{l=k}^{k+\tau} h^{f}_{\mathbf{d},i}(l)} u^{\star}_{\mathbf{sh,u},i}(m), \tag{3-44}$$

$$\bar{\boldsymbol{u}}_{\mathbf{sc},i}(k:k+\tau) = \frac{\boldsymbol{c}^{f}_{\mathbf{d},i}(k:k+\tau)}{\sum_{l=k}^{k+\tau} c^{f}_{\mathbf{d},i}(l)} u^{\star}_{\mathbf{sc,u},i}(m), \tag{3-45}$$

where $h^{f}_{\mathbf{d},i}$ indicates the demand forecast. So, $u^{\star}_{\mathbf{sh,u},i}(m)$ and $u^{\star}_{\mathbf{sc,u},i}(m)$ are divided proportionally to the forecast of the demand. The advantage of having this formulation is that if the demand forecast is accurate enough, we have an appropriate distribution of the usage of the seasonal storage.

Furthermore, it is also important to note that the upper layer has to send the upper bounds for 2τ time steps in which $\bar{\boldsymbol{u}}_{\mathbf{sh},i}(k+\tau+1:k+2\tau)$ and $\bar{\boldsymbol{u}}_{\mathbf{sc},i}(k+\tau+1:k+2\tau)$ are derived from Equation (3-44) and (3-45) by using $u^{\star}_{\mathbf{sh,u},i}(m+1)$, $u^{\star}_{\mathbf{sc,u},i}(m+1)$, $\boldsymbol{h}^{f}_{\mathbf{d},i}(k+\tau+1:k+2\tau)$ and $\boldsymbol{c}^{f}_{\mathbf{d},i}(k+\tau+1:k+2\tau)$. It is because the first N_h values of $\bar{\boldsymbol{u}}_{\mathbf{sh},i}(k+\tau+1:k+2\tau)$ and $\bar{\boldsymbol{u}}_{\mathbf{sc},i}(k+\tau+1:k+2\tau)$ are required in the lower layer.

3-4-3 Dealing with Uncertainties

The randomization-based method is again considered to deal with the chance constraints that appear in both layers of the hierarchical MPC. The optimization problem of the upper layer is convex since the cost function and the constraints are convex, so the scenario approach can be applied to it. So, the chance constraint (3-38b) is approximated with a number of deterministic constraints as the following:

$$\mathbf{A}\boldsymbol{x}_{\mathbf{u},i}(m) + \mathbf{B}_{\mathbf{u},i}\tilde{\boldsymbol{u}}_{\mathbf{u},i} + \mathbf{H}_i\tilde{\boldsymbol{w}}^{(s)}_{\mathbf{u},i} \geq 0, \quad \forall i \in \mathcal{N} \quad \forall s \in \{1, \cdots, N_{s,\mathbf{u}}\}, \tag{3-46}$$

where the number of scenarios, $N_{s,\mathbf{u}}$ depends on the number of the decision variables in this problem, $N_{d,\mathbf{u}} = 5NN_{h,\mathbf{u}}$ as

$$N_{s,\mathbf{u}} \geq \frac{2}{\varepsilon}\left(N_{d,\mathbf{u}} + \ln\frac{1}{\beta}\right), \tag{3-47}$$

in which β is the level of confidence. Meanwhile, the robust randomized method is employed in the lower layer since the problem is mixed-integer which is not convex.

3-4-4 Distributed MPC in the Upper Layer

A fully distributed MPC scheme can be achieved by decomposing the upper layer problem since the lower layer problem is already separable. Since by definition the cost function is strictly convex, the standard dual decomposition can be applied to this problem. Thus by assigning the Lagrange Multipliers $\boldsymbol{\lambda}_i \in \mathbb{R}^{2N_{h,\mathbf{u}}}$ to the coupling constraints (3-38e), we obtain the corresponding Lagrangian as

$$L = \sum_{i=1}^{N} \boldsymbol{J}_{\mathbf{u},i}(\tilde{\boldsymbol{u}}_{\mathbf{u},i}) + \boldsymbol{\lambda}_i^T\left(\mathbf{G}_{ii}\left(\boldsymbol{\Phi}_i\boldsymbol{x}_{\mathbf{a},\mathbf{u},i}(m) + \boldsymbol{\Gamma}_i\tilde{\boldsymbol{u}}_{\mathbf{u},i}\right) + \sum_{j\in\mathcal{N}_{-i}}\mathbf{G}_{ij}\left(\boldsymbol{\Phi}_j\boldsymbol{x}_{\mathbf{a},\mathbf{u},j}(m) + \boldsymbol{\Gamma}_{s,j}\tilde{\boldsymbol{u}}_{\mathbf{u},j}\right) - \boldsymbol{\Theta}_i\right) \tag{3-48}$$

which can be arranged into

$$L = \sum_{i=1}^{N}\left(\boldsymbol{J}_{\mathbf{u},i}(\tilde{\boldsymbol{u}}_{\mathbf{u},i}) + \left(\boldsymbol{\lambda}_i^T\mathbf{G}_{ii} + \sum_{j\in\mathcal{N}_{-i}}\boldsymbol{\lambda}_j^T\mathbf{G}_{ji}\right)\left(\boldsymbol{\Phi}_i\boldsymbol{x}_{\mathbf{a},\mathbf{u},i}(m) + \boldsymbol{\Gamma}_i\tilde{\boldsymbol{u}}_{\mathbf{u},i}\right) - \boldsymbol{\lambda}_i^T\boldsymbol{\Theta}_i\right). \tag{3-49}$$

The Lagrangian (3-49) is separable. Moreover, by omitting the constant terms, we obtain N local optimization problems which are assigned to the agents as follows:

$$\begin{aligned} \underset{\tilde{\boldsymbol{u}}_{\mathbf{u},i}}{\text{minimize}} \quad & \boldsymbol{J}_{\mathbf{u},i}(\tilde{\boldsymbol{u}}_{\mathbf{u},i}) + \left(\boldsymbol{\lambda}_i^T\mathbf{G}_{ii} + \sum_{j\in\mathcal{N}_{-i}}\boldsymbol{\lambda}_j^T\mathbf{G}_{ji}\right)\boldsymbol{\Gamma}_i\tilde{\boldsymbol{u}}_{\mathbf{u},i} \\ \text{subject to} \quad & \text{(3-38b), (3-38c), (3-38d), (3-38f).} \end{aligned} \tag{3-50}$$

The dual ascent method is then can be applied which means that Problem (3-50) is solved iteratively with certain $\boldsymbol{\lambda}_i^{(q)} \quad \forall i \in \mathcal{N}$ where (q) indicates the iteration step. From the solutions of Problem (3-50), $\tilde{\boldsymbol{u}}^{(q)}_{\mathbf{u},i}$ $\forall i \in \mathcal{N}$, the Lagrange Multipliers $\boldsymbol{\lambda}_i$ $\forall i \in \mathcal{N}$ are updated as follows:

$$\boldsymbol{\lambda}_i^{(q+1)} = \max\left(0, \boldsymbol{\lambda}^{(q)} + \frac{\alpha_g}{r}\boldsymbol{d}_i\right), \tag{3-51}$$

in which the directions of the ascent method d_i $\forall i \in \mathcal{N}$ are computed as

$$\boldsymbol{d}_i = \mathbf{G}_{ii}\left(\boldsymbol{\Phi}_i \boldsymbol{x}_{\mathbf{a},\mathbf{u},i}(m) + \boldsymbol{\Gamma}_i \tilde{\boldsymbol{u}}_{\mathbf{u},i}^{(q)}\right) + \sum_{j \in \mathcal{N}_{-i}} \mathbf{G}_{ij}\left(\boldsymbol{\Phi}_j \boldsymbol{x}_{\mathbf{a},\mathbf{u},j}(m) + \boldsymbol{\Gamma}_{s,j} \tilde{\boldsymbol{u}}_{\mathbf{u},j}^{(q)}\right) - \boldsymbol{\Theta}_i. \tag{3-52}$$

Note that the step size is diminishing with $\alpha_g \in [0, 1]$ and due to the fact that the Lagrange multipliers are assigned to the inequality constraints, they are non-negative at each iteration. The stopping criterion of this iterative procedure comes from the complementary slackness condition due to the fact that the coupling constraints are inequality constraints. Hence, the stopping criterion is

$$\boldsymbol{\lambda}_i^{(q)T} \boldsymbol{d}_i < \epsilon \quad \forall i \in \mathcal{N}, \tag{3-53}$$

for a small positive ϵ.

The proposed formulation allows the upper layer problem to be solved in a fully distributed setting in which some information need to be exchanged between the agents, such as the Lagrange multipliers in order to solve the local optimization problem and the prediction of the states of the seasonal storages in order to update the Lagrange multipliers. It is also possible to assign a coordinator to update the Lagrange multipliers and in this setting the information is exchanged between the coordinator and the agents. The complete procedure of the proposed hierarchical MPC is given in Algorithm 4.

Algorithm 4 Hierarchical MPC

for $k = 1, 2, \ldots$ **do**
  if mod $(k, \tau) = 1$ **then**
    (Solving the upper layer problem (3-38) in a distributed fashion)
    $m = \lceil k/\tau \rceil$.
    Initialize $\tilde{\boldsymbol{\lambda}}_i^{(1)}$, $\forall i \in \mathcal{N}$, $\rho > 0$ randomly if $m = 1$, otherwise using warm start.
    q=1.
    while Equation (3-53) is false **do**
      Each agent sends $\tilde{\boldsymbol{\lambda}}_i^{(q)}$ to its neighbors.
      Each agent solves Problem (3-50) using the scenario approach
      to obtain $\tilde{\boldsymbol{u}}_{\mathbf{u},i}^{(q)}$.
      Each agent sends $\tilde{\boldsymbol{u}}_{\mathbf{u},i}^{(q)}$ to its neighbors.
      Each agent updates $\tilde{\boldsymbol{\lambda}}_i$ using Equation (3-51).
      $q = q + 1$.
    end while
    Compute $\bar{\boldsymbol{u}}_{\mathbf{sh},i}(k : k + 2\tau)$ and $\bar{\boldsymbol{u}}_{\mathbf{sc},i}(k : k + 2\tau)$ using Equation (3-44) and (3-45)
    Send $\bar{\boldsymbol{u}}_{\mathbf{sh},i}[k : k + 2\tau]$ and $\bar{\boldsymbol{u}}_{\mathbf{sc},i}[k : k + 2\tau]$ to the lower layer MPC.
  end if
  Using the robust randomized method, each agent solves the lower layer problem:

$$\begin{aligned} \underset{\{\tilde{\boldsymbol{u}}_i, \tilde{\boldsymbol{v}}_i\}_{i=1}^N}{\text{minimize}} \quad & \boldsymbol{J}_i(\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{u}}_i) \\ \text{subject to} \quad & \text{(3-35b), (3-35c), (3-35d), (3-35e), (3-40), (3-41).} \end{aligned} \tag{3-54}$$

  Each agent apply $\tilde{\boldsymbol{u}}_i^\star(k)$, $\tilde{\boldsymbol{v}}_i^\star(k)$.
end for

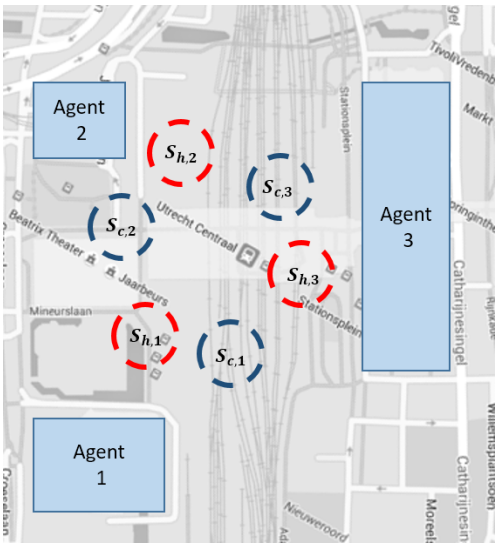


Figure 3-4: A 3-agent thermal grid in Utrecht.

3-5 Numerical Study of the Second Case

As a numerical study, we consider a realistic three-agent thermal grid which is located in the city center of Utrecht as shown in Figure 3-4. The agents in the grid are 3 office buildings, each of which owns a seasonal storage that is assumed to have dynamic as described in 3-2-1. The heating and cooling demands of all agents, which were computed via LEA model based on the actual buildings' sizes (Table A-3) and the weather data are shown in Figure 3-5 until 3-7. Moreover, we also assume that the amount of supplies that are stored in the storages can be measured and the positions of the storages are according to Figure 3-4. Note that the red and blue circles in Figure 3-4 represent the warm and cold wells respectively.

Based on the assumption of the location of the wells, we can infer the coupling constraints between agent as

$$\begin{aligned} V_{\mathbf{h},1}(k) + V_{\mathbf{c},2}(k) &\leq \bar{V}_{\mathbf{hc},12}, \\ V_{\mathbf{h},2}(k) + V_{\mathbf{c},3}(k) &\leq \bar{V}_{\mathbf{hc},23}, \\ V_{\mathbf{h},3}(k) + V_{\mathbf{c},1}(k) &\leq \bar{V}_{\mathbf{hc},31}, \end{aligned}$$

where $\bar{V}_{\mathbf{hc},12} = \bar{V}_{\mathbf{hc},23} = \bar{V}_{\mathbf{hc},31} = 2.4 \cdot 10^5$ m^3 which was chosen by considering the maximum possible volume of one of the wells. Furthermore, the values of the parameters of each agent are given in Table 3-1. The initial conditions of the cold storages $s_{\mathbf{c},i}[0]\ \forall i \in \mathcal{N}$ were set to be approximately 60% of the total cooling demand so that the storage could not fulfill the all the demand and the local production units (the chillers) had to be used. On the other hand, the initial supply in the hot storages $s_{\mathbf{h},i}[0]\ \forall i \in \mathcal{N}$ were set to be zero since the simulation started in the summer season which mostly requires cooling supplies. Additionally, the temperature of the wells were assumed to be constant and the temperature difference between one well and the ambient aquifer is 4K which yields the values of $\alpha_{\mathbf{h}}$ and $\alpha_{\mathbf{c}}$ as in Table 3-1. This assumption was made by considering the possible temperature ranges of the wells and the

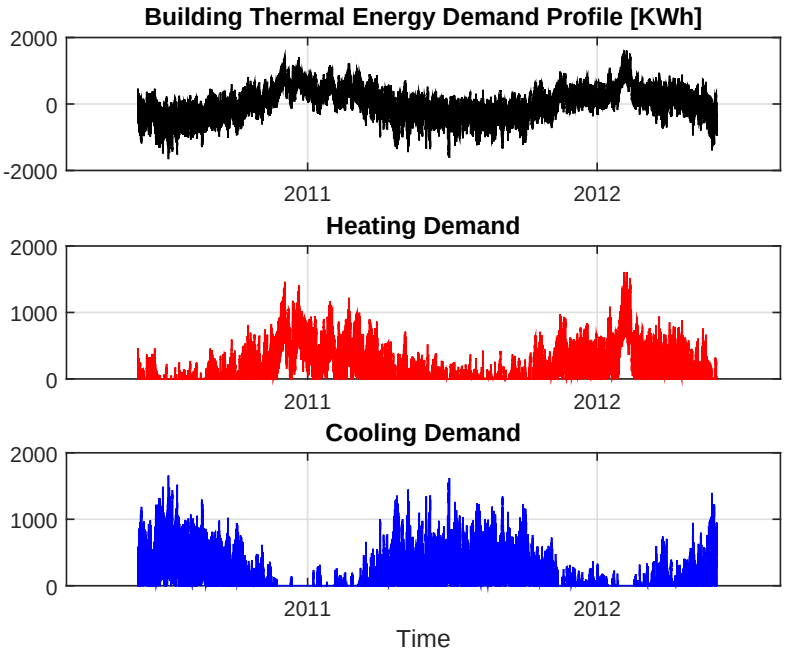


Figure 3-5: Heating and cooling demand of agent 1.

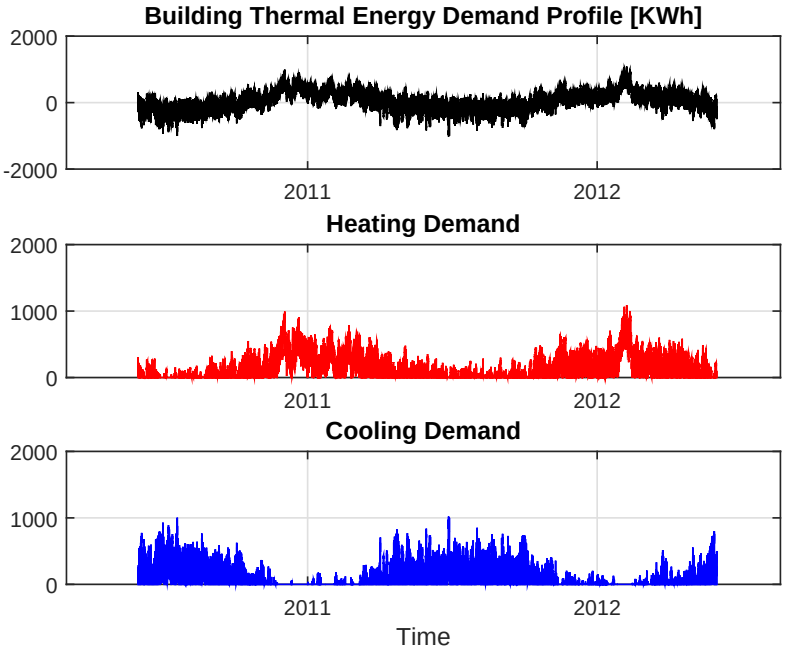


Figure 3-6: Heating and cooling demand of agent 2.

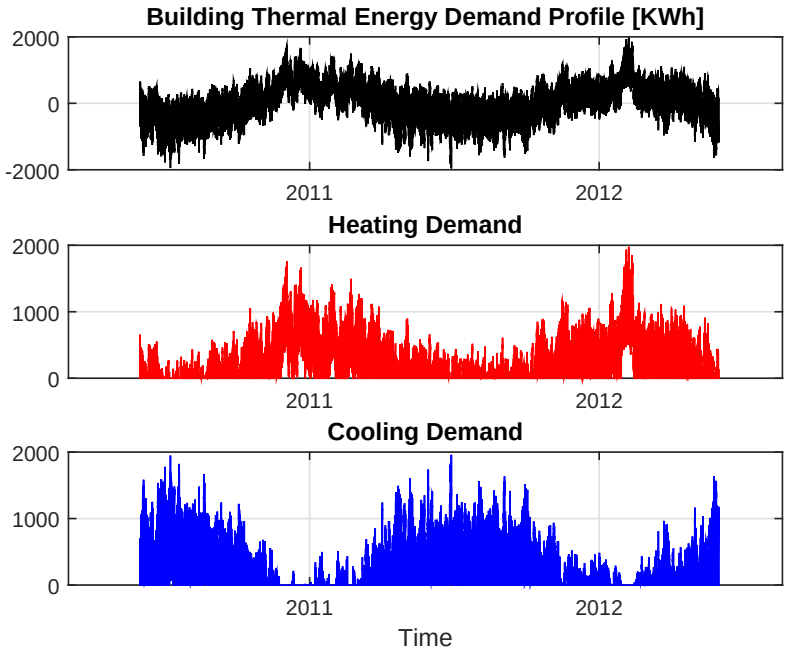


Figure 3-7: Heating and cooling demand of agent 3.

Table 3-1: The Values of the Parameters

Parameter	Agent 1	Agent 2	Agent 3	Unit
$h_{\mathbf{b}}^{\min}$, $h_{\mathbf{b}}^{\max}$, $c_{\mathbf{ch}}^{\min}$,$c_{\mathbf{ch}}^{\max}$	10, 2000, 10, 2000	10, 2000, 10, 2000	10, 2000, 10, 2000	KWh
$h_{\mathbf{im}}^{\min}$,$h_{\mathbf{im}}^{\max}$, $c_{\mathbf{im}}^{\min}$,$c_{\mathbf{im}}^{\max}$	0, 10000, 0, 10000	0, 10000, 0, 10000	0, 10000, 0, 10000	KWh
$u_{\mathbf{sh}}^{\min}$, $u_{\mathbf{sh}}^{\max}$, $u_{\mathbf{sc}}^{\min}$, $u_{\mathbf{sc}}^{\max}$	0, 2000, 0, 2000	0, 2000, 0, 2000	0, 2000, 0, 2000	KWh
$s_{\mathbf{h}}[0], s_{\mathbf{c}}[0]$,$\boldsymbol{x}[0]$, $\bar{s}_{\mathbf{hc}}$	0, $8\cdot10^5$, $\mathbf{0}$, $8\cdot10^5$	0, $4.8\cdot10^5$, $\mathbf{0}$, $4.8\cdot10^5$	0, $9.6\cdot10^5$, $\mathbf{0}$, $9.6\cdot10^5$	KWh
$\eta_{\mathbf{buf}}$, $\alpha_{\mathbf{hp}}$	0.9, 1.35	0.9, 1.35	0.9, 1.35	-
$\alpha_{\mathbf{h}}$, $\alpha_{\mathbf{c}}$	4.67, 4.67	4.67, 4.67	4.67, 4.67	KWh/m^3
Q, $r_{\mathbf{b}}$, $r_{\mathbf{ch}}$, $r_{\mathbf{im}}$	100, 10, 10, 250	100, 10, 10, 250	100, 10, 10, 250	-
$r_{\mathbf{ush}}$, $r_{\mathbf{usc}}$, $\Lambda_{\mathbf{b}}^{\mathrm{su}}$,$\Lambda_{\mathbf{ch}}^{\mathrm{su}}$	3, 3, 1, 1	3, 3, 1, 1	3, 3, 1, 1	-

ambient aquifers. It is also worth to mention that the weight of the supply of the seasonal storages in the cost function were set to be smaller than the weight of the other supplies (the boilers and the imported energy from the external party).

In this one-year simulation, we investigated two control approaches, the centralized approach which means the MPC controller solves Problem 3-35 with small prediction horizon of 24 hours at each sampling time and the hierarchical approach as written in Algorithm 4 in which the prediction horizon of the upper layer was $m = 12$ months and the lower layer was 24 hours. The simulations were done in MATLAB by using Gurobi [39] as the solver of the optimizations and YALMIP [38] as an interface. Figure 3-8 until 3-13 depict the results of the simulations.

The evolutions of the water volume in the warm and cold wells of the ATES are depicted in Figure 3-8. When the system was controlled by the hierarchical MPC scheme, the growing and reducing rate of the volume of the wells were slightly slower. This happened due to

the fact the volume of the water that can be extracted and injected were limited based on the decisions of the upper layer optimization. Moreover, as seen in Figure 3-9 that both controllers were able to provide solutions which were feasible for the coupling constraints. As mentioned earlier that these constraints are important to maintain the efficiency and the sustainability of the ATES.

Figure 3-10 and 3-11 show the usages of the boilers and chillers in the system which can be considered to be minimal. The peak usage of the boilers happened between February and April while the chillers worked the hardest between September and October. This is interesting since they occurrence did not match with the peak of the heating and cooling demands which happened earlier. However, this result is logical since during those periods the seasonal storage could not provide heating and cooling supply. As seen in Figure 3-8, the warm wells were completely depleted from February until April while the cold wells were empty[1] in September and October. Therefore, it can be said that in the proposed energy management the boilers and chillers were considered as the backups when the seasonal storages could not provide the supplies.

When comparing the decision of the usage of the boilers and chillers of both approach, we can see in Figure 3-10 and 3-11 that the centralized MPC approach tended to not use the production units while the seasonal storages were still able to provide the supplies and to maximize the usage of them when the seasonal storages did not have supplies anymore. Meanwhile the hierarchical MPC approach tended to distribute the load, even though with a small proportion, to the production units during the time when the seasonal storages only had limited supplies, *i.e.* they were almost depleted. These can be seen in the period of January-February for the plots of the boilers (Figure 3-10) and in the period of June-July for the plots of the chillers (Figure 3-11). Thus resulted in shorter period of the depletion of the wells as seen in Figure 3-8.

Even though the total cost of the hierarchical MPC scheme was more expensive, the difference, which was approximately 5%, was considerably small (see Figure 3-12). It is argued that it was caused by the errors of the demand forecast which were used as the stochastic vectors. In the upper layer, since the decision of the monthly supply usage of the storages were distributed proportionally with the ratio of the current and the total demand of that month, the differences between the forecast and the actual demand imply mismatches on the distribution of the upper bounds. It sometimes resulted in the usage of the production units which were more costly as can be seen as the small spikes that appear in Figure 3-10 and 3-11. Preliminary simulations which had been carried out and in which both controllers had the perfect demand forecast and did not apply robust randomized method show that the hierarchical MPC could provide cheaper cost up to 20%. Therefore, having accurate demand forecasts is important to improve the cost of the hierarchical MPC.

At most of the time, both controllers provided positive imbalance errors. Figure 3-13 shows the imbalance errors of agent 1 in the simulation of the hierarchical MPC. However, there existed some violations on the non-negativity constraints that were imposed on, which were expected, due to the fact that we imposed probabilistic constraints which allow some small fraction of violations. Moreover, the violations (negative imbalance) happened at most 0.4% of the total simulation duration which was smaller than the level of violation that was predefined

[1] Physically, the wells were not empty, but the temperature differences between the water in the wells and the ambient aquifer were zero which means the wells did not have energy supply.

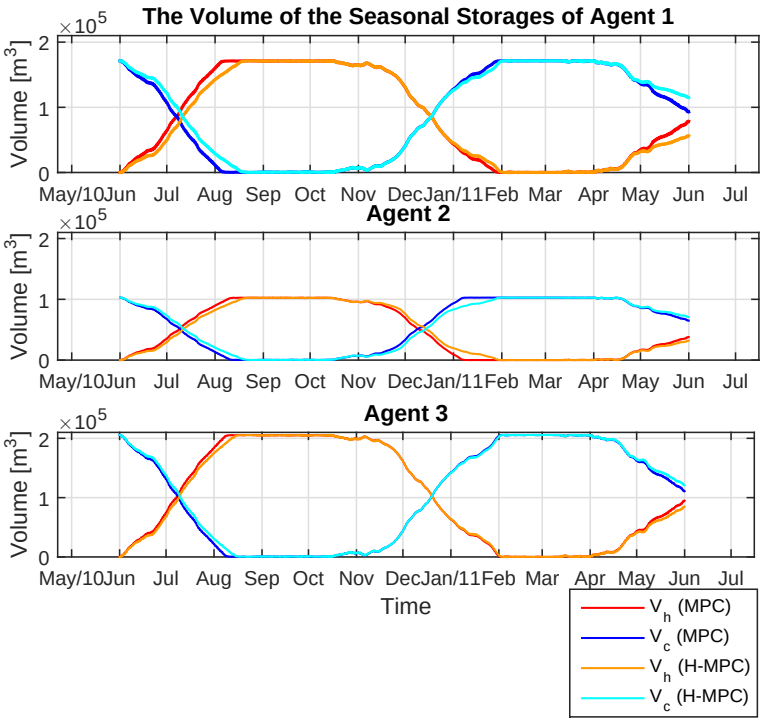


Figure 3-8: The evolution of the supply level of the seasonal storages.

as 10%. Hence it demonstrates that the robust randomized approach that was applied was able to satisfy the given probabilistic constraints.

In conclusion, the simulation results show that the proposed controllers were able to manage the thermal grid to be reliable by ensuring that the demand were mostly satisfied. Moreover, the coupling constraints that were imposed to maintain the sustainability and the efficiency of the ATES were also satisfied when the system was controlled by both approaches. Furthermore, we also have shown that the hierarchical MPC approach is promising to be implemented since it was able to provide an intelligent way of limiting the usage of the supply of the seasonal storage and a distributed computation scheme.

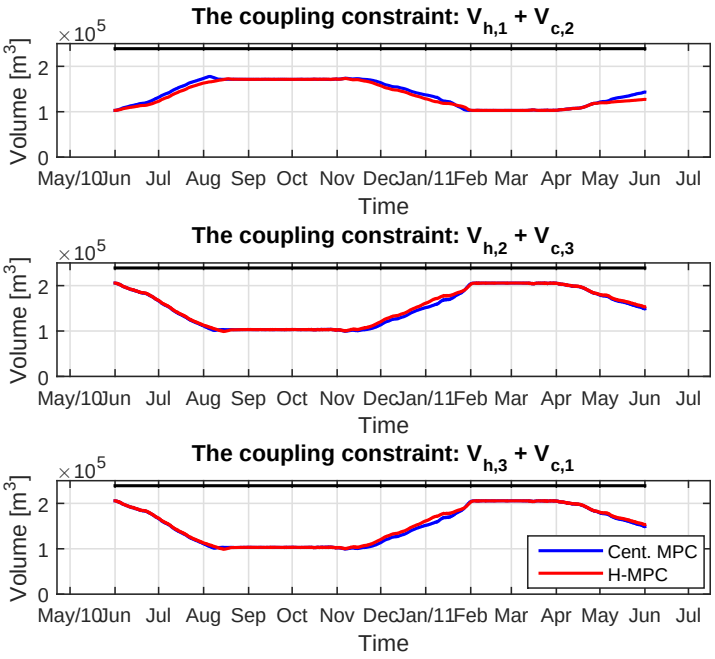


Figure 3-9: The coupling constraints were satisfied.

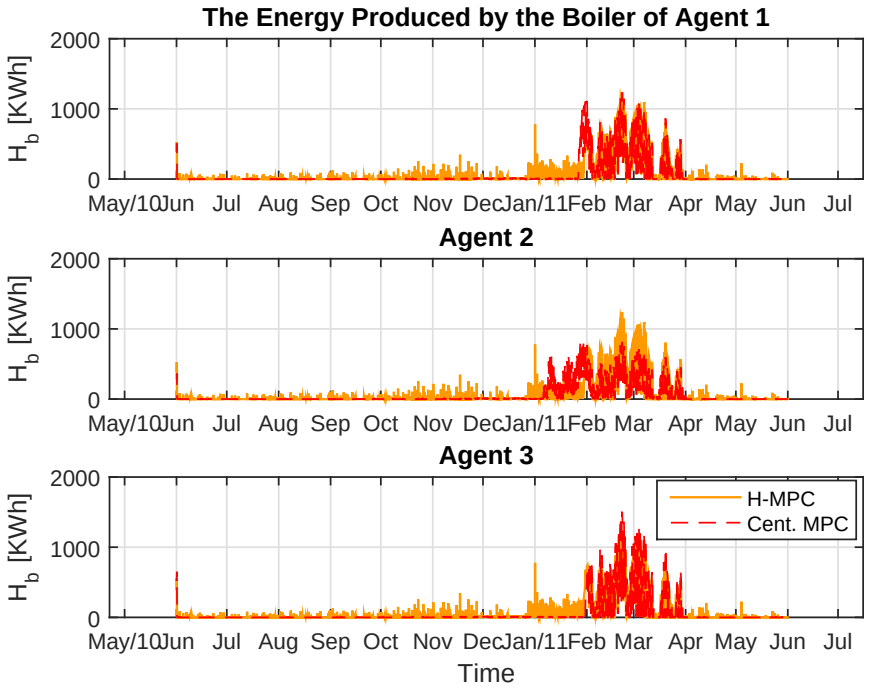


Figure 3-10: The usage of the boilers.

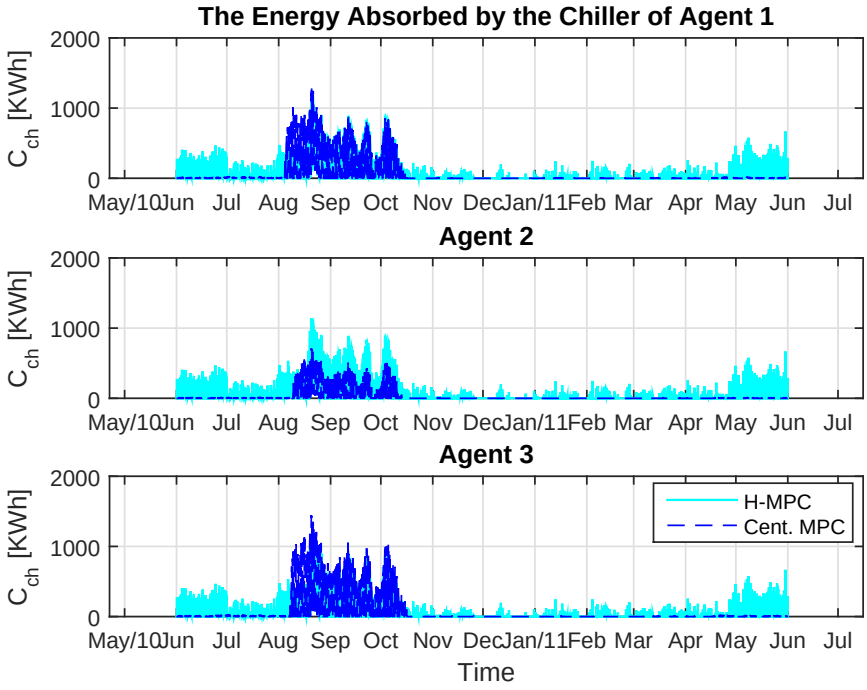


Figure 3-11: The usage of the chillers.

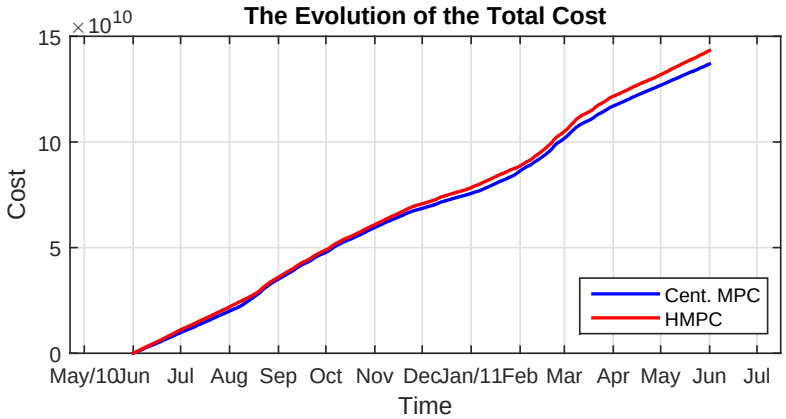


Figure 3-12: The accumulation of the total cost of all agents.

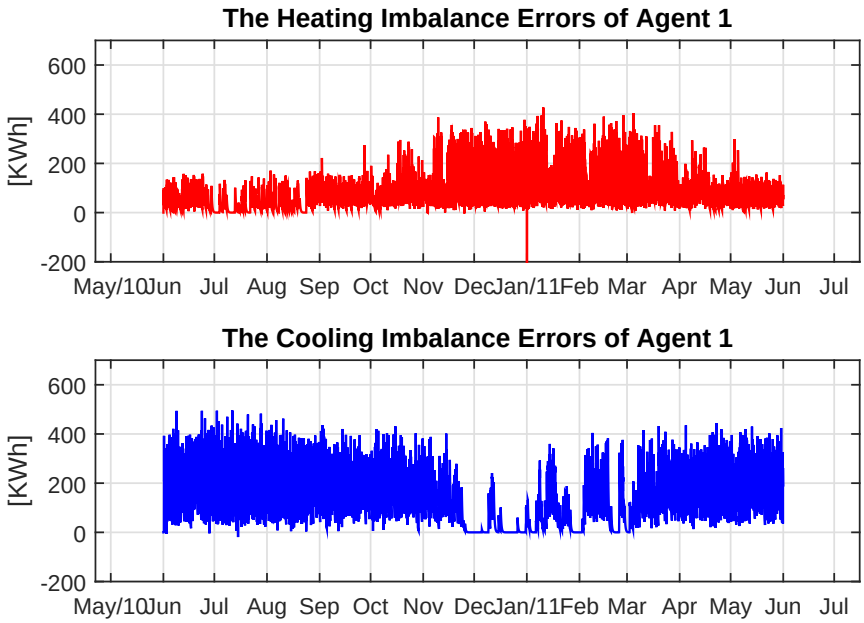


Figure 3-13: The imbalance errors.

Chapter 4

Conclusions and Future Work

The focus of this research is to develop distributed Model Predictive Control (MPC) approaches which act as the energy management systems of thermal grids. In this chapter we provide some concluding remarks regarding the work that has been done and some recommendations for the future work which are related to this project.

4-1 Conclusions

1. In this thesis, two cases of thermal grids have been investigated. The main similarity of them, which is the main focus of our work, is the existence of the coupling between agents which makes the problem not trivially separable. Thus, for each case, we have proposed distributed approaches to solve the mixed-integer chance constrained optimization problem of production planning.

2. In the first case in which the thermal grid is intended to fulfill heating demands of the agents by having decentralized production units such as boilers and micro-CHP and the possibility of exchanging energy between agents, the proposed distributed robust randomized methods which are based on the Alternating Direction Method of Multipliers (ADMM) were able to have similar performance in term of the cost as the centralized counterpart in the simulation. This means that by having distributed optimization methods we are able to add some advantages such as lower computational demands, scalability and flexibility to the energy management system which is proposed in [1] without losing the performance.

3. In the second case in which the grid have seasonal storages, the hierarchical MPC is proposed not only to provide a distributed computation method but also to capture the dynamic of seasonal storages in a long period of time. The dual ascent method is applied to decompose the problem in order to obtain a distributed scheme. Simulation results showed that it was able to ensure the grid to satisfy all constraints including the coupling between the agents. Additionally, it also provides an intelligent method to limit the usage of the seasonal storages.

4-2 Future Work

Reflecting on the work presented on this thesis, some recommendations for future work are stated in this section.

1. The distributed robust randomized methods which are proposed in Chapter 2 have considerably low convergence rate. An improvement on this area may be achieved by applying a fast or accelerated gradient method when updating the Lagrange multipliers and investigating structured methods to set the penalty term ρ.

2. The distributed energy management in a thermal grid requires the agents to share or exchange information. In this regard, some scenarios such as delays on exchanging information or the resistance of some agents to share the information need to be investigated and tackled.

3. In the second case study, the dynamics of the seasonal storages in this thesis is modeled as a linear discrete-time system. However, these storages, such as Aquifer Thermal Energy Storage (ATES) have non-linear behavior, as also noted in [43]. Incorporating the nonlinearities of the dynamics of the seasonal storage yields to a more accurate model on the controller. Thus it is expected that the controller will provide better decisions.

4. Physically, the coupling constraints on the volume of the wells are related with the distance between those two wells. We may consider to replace the linear volume-based coupling constraints with the coupling of the radius of the wells. However, these constraints are not linear with respect to the states of the seasonal storage. Therefore, methods to linearly approximate the nonlinear constraints need to be investigated so that the complexity of the problem does not increase.

Appendix A

Mathematical Model of the Thermal Demands

In this part we explain the mathematical model that is used to generate the thermal energy demand of an agent and provide an example. Section A-1 provides the explanation of the model as well as all equations of the model. In Section A-2, we show the example of the daily thermal demand profile of a building.

A-1 Low Energy Architecture Model

The model of the thermal energy demand of an agent (building or house) is obtained according to the Low Energy Architecture (LEA) model which is based on the following dynamics of the indoor air temperature [35]:

$$\frac{dT_{\text{in}}}{dt} c_{\text{p,a}} \rho_{\text{a}} V_{\text{a}} = h_{\text{d}} + h_{\text{int}} + h_{\text{env}}, \tag{A-1}$$

where h_{d} [KWh] is the thermal energy demand and h_{int} [KWh] is the thermal energy that is produced internally, by the occupancy of people, lighting, and other installation while h_{env} [KWh] is the thermal energy gain from/loss to the environment. Moreover, T_{in} [K] is the desired indoor air temperature, c_{pa} [J/kgK] is the specific heat capacity of air, ρ_{a} [kg/m^3] is the air density, and V_{a} [m^3] is the volume of the air inside the building. In the equilibrium point, that is $dT_{\text{in}} = 0$, where we assume that the desired indoor air temperature has been achieved and is kept constant, we have

$$h_d = -h_{\text{int}} - h_{\text{env}}(T_{\text{in}}, T_{\text{out}}, T_{\text{soil}}, v_{\text{w}}, q_{\text{sol}}),$$

where T_{out} [K] and T_{soil} [K] denote the outside air and soil temperature, v_{w} [m/s] is the wind speed, and q_{sol} [KWh] is the solar radiation. These variables are related to the environment conditions and vary over time.

The internal energy, h_{int} is the heat that is produced due to the occupancy of the building. It is calculated as follows:

$$h_{\text{int}} = h_{\text{person}} + h_{\text{inst}} + h_{\text{light}}, \tag{A-2}$$

where h_{person} is the heat that is transmitted by people. As in [35], the heat that is produced per person is assumed to be constant, that is $C_{h_{\text{person}}} = 80$ [KWh]. So, h_{person} depends on the number of people (N_{person}), as

$$h_{\text{person}} = C_{h_{\text{person}}} N_{\text{person}} c_{f,\text{person}}, \tag{A-3}$$

where $c_{f,\text{per}}$ is the covection factor of a person. Moreover, h_{inst} and h_{light} are the heat that are produced by the installations and the lighting, respectively. They are assumed to be constant per area, *i.e.* $C_{h_{\text{inst}}} = C_{h_{\text{light}}} = 10$ [J/m^2] [35], thus for the total area is represented as

$$h_{\text{inst}} = C_{h_{\text{inst}}} a_{\text{floor}} c_{f,\text{inst}}, \tag{A-4}$$

$$h_{\text{light}} = C_{h_{\text{light}}} a_{\text{floor}} c_{f,\text{light}}, \tag{A-5}$$

where a_{floor} is the total floor area of the building and $c_{f,\text{inst}}, c_{f,\text{light}}$ are the convection factor of the installation and the lighting, respectively.

The thermal energy that is lost to or gained from the environment, h_{env}, is broken down into three parts as follows:

$$h_{\text{env}} = h_{\text{soil}} + h_{\text{shell}} + h_{c,\text{solar}}. \tag{A-6}$$

The heat that is transferred to/from the soil, h_{soil}, which is caused by the temperature difference between the indoor air and the soil is represented by:

$$h_{\text{soil}} = (T_{\text{soil}} - T_{\text{in}}) U_{\text{bot}} a_{\text{soil}}, \tag{A-7}$$

where U_{bot} [J/m^{2}K] is the overall heat transfer coefficient to the soil, and a_{soil} [m^2] is the total surface area of the building that touches the ground. Whereas, the heat exchanged which is due to the temperature difference between the indoor and the outdoor air and transfered via the facade is denoted by h_{shell} and formulated as follows

$$h_{\text{shell}} = (T_{\text{out}} - T_{\text{in}}) \sum_{i=1}^{n_{\text{fac}}} (U_{\text{wall}} a_{\text{wall}} + U_{\text{win}} a_{\text{win}}), \tag{A-8}$$

where U_{wall} [J/m^{2}K] and U_{win} [J/m^{2}K] are the overall heat transfer coefficient of the walls and windows, respectively, and a_{wall} [m^2] and a_{win} [m^2] are the surface area of the walls and windows respectively. The heat transfer coefficient U_{wall} varies depending on the wind speed as follows [35]:

$$U_{\text{wall}} = \frac{1}{(\frac{1}{h_{c,\text{in}}} + R_{c,\text{wall}} + \frac{1}{h_{c,\text{out}}})}, \tag{A-9}$$

where $h_{c,\text{in}}$ [J/m^{2}K] is the heat transfer coefficient at the inside of the walls, $R_{c,\text{wall}}$ [m^{2}K/J] is the conductive heat transfer resistance, and $h_{c,\text{out}}$ [J/m^{2}K] denotes the heat transfer coefficient at the outside of the walls which depends on the wind speed as follows:

$$h_{c,\text{out}} = \begin{cases} 1.2 v_{\text{w}} + 3.8 & \text{if } v_{\text{w}} < 4 \\ 2.5 v_{\text{w}}^{0.8} & \text{otherwise} \end{cases} \tag{A-10}$$

Moreover, the solar radiation, $h_{c,\text{solar}}$, is considered to be transfered as convective thermal energy to the building. This is computed as sum of thermal energy from solar that is transferred via the windows of all facades as follows:

$$h_{c,\text{solar}} = \sum_{i=1}^{n_{\text{fac}}} \left(SGF_i \ \gamma_{\text{w}} \ a_{\text{shell},i} \ c_{f,\text{win}} \ q_{\text{sol}} \right), \tag{A-11}$$

where SGF is the solar gain factor, γ_{w} is the window fraction, $a_{\text{shell},i}[\text{m}^2]$ is the total surface area of the shell, *i.e* $a_{\text{shell},i} = a_{\text{wall},i} + a_{\text{win},i}$, c_f is the convection factor of the window $[\text{J/m}^2]$.

A-2 Generating Thermal Demand Profile

The example of the daily thermal energy demand profile of a building which is based on the LEA model is provided in Figure 2-1 which is again showed as the red plot in the following Figure A-1. It was generated using a 24-hour weather data ($T_{\text{in}}, T_{\text{out}}, T_{\text{soil}}, v_{\text{w}}, q_{\text{sol}}$) as depicted in Figure A-2 and the parameters as in Table A-1. In addition, in all our numerical studies, we used the same values of the convection factors and the heat transfer coefficients as given in Table A-1 while the building parameters were different between agents. It is assumed that the surface area at each floor is equal . Moreover, the area of window a_{win} and the area of the wall a_{wall} can be obtained from γ_{w} and a_{shell}. Therefore, the total area of the floor as well as the total area of the windows and the walls are calculated as follows:

$$\begin{aligned} a_{\text{floor}} &= n_{\text{floor}} \ a_{\text{s,floor}}, \\ a_{\text{win}} &= \gamma_{\text{w}} \ a_{\text{shell}}, \\ a_{\text{wall}} &= (1 - \gamma_{\text{w}}) a_{\text{shell}}. \end{aligned}$$

Furthermore, we assume that all agents were located close to each other so the same weather data applies to them.

Table A-1: The Parameters of the LEA Model

Building Parameters	Convection Factors	Heat Transfer Coefficients
$a_{\text{s,floor}} = 480\text{m}^2$	$c_{f,\text{win}} = 0.024$	$U_{\text{bot}} = 1.3$ $[\text{J/m}^2\text{K}]$
$n_{\text{floor}} = 2$	$c_{f,\text{per}} = 0.02$	$U_{\text{win}} = 1.6$ $[\text{J/m}^2\text{K}]$
$a_{\text{shell}} = 840\text{m}^2$	$c_{f,\text{light}} = 0.08$	$h_{c,\text{in}} = 8$ $[\text{J/m}^2\text{K}]$
$\gamma_{\text{w}} = 0.5$	$c_{f,\text{ins}} = 0.23$	$R_{c,\text{wall}} = 2.5$ $[\text{m}^2\text{K/J}]$
$SGF = 0.5$		

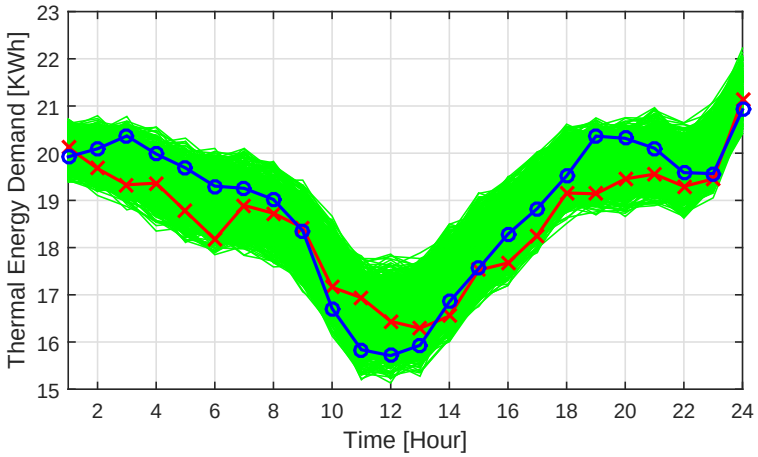


Figure A-1: Red line: the real thermal energy demand generated using the LEA model.

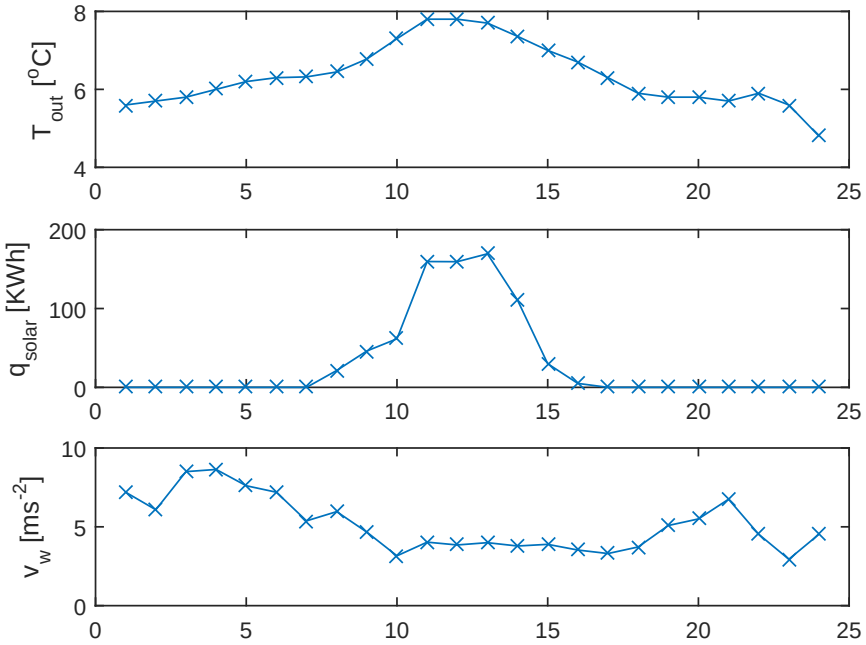


Figure A-2: The 24-hour weather data that was used to generate the demand profile in Figure A-1.

A-3 Building Parameters of the Case Studies

Table A-2 and A-3 show the values of the building parameters that are necessary to generate the thermal energy demands and its forecasts in each case study.

Table A-2: The Building Parameters of the First Case Study (Chapter 2)

Building	$a_{\text{s,floor}}$ [m^2]	a_{shell} [m^2]	n_{floor}	γ_{w}	$T_{\text{in}}[^o\text{C}]$
Agent 1	500	1200	3	0.4	20
Agent 2	480	840	2	0.5	20
Agent 3	460	720	3	0.3	20

Table A-3: The Building Parameters of the Second Case Study (Chapter 3)

Building	$a_{\text{s,floor}}$ [m^2]	a_{shell} [m^2]	n_{floor}	γ_{w}	$T_{\text{in}}[^o\text{C}]$
Agent 1, sub building 1	4275	2664	4	0.95	18
Agent 1, sub building 2	4275	5328	8	0.95	18
Agent 1, sub building 3	2513	20991	29	0.95	18
Agent 2	19125	14000	4	0.4	19
Agent 3, sub building 1	2909	16612	22	0.6	20
Agent 3, sub building 2	2909	16612	22	0.6	20
Agent 3, sub building 3	2778	6641	9	0.5	20
Agent 3, sub building 4	3889	7875	9	0.5	20

Appendix B

Formulas and Detailed Definitions

This appendix provides the definitions of the compact vectors and matrices.

B-1 Dynamics Equations

The state and disturbance predictions as well as the decision variables which are written in compact formats are defined as the following

$$\tilde{\boldsymbol{x}}_i = \begin{bmatrix} \boldsymbol{x}_i[k+1] \\ \boldsymbol{x}_i[k+2] \\ \vdots \\ \boldsymbol{x}_i[k+N_h] \end{bmatrix}, \tilde{\boldsymbol{u}}_i = \begin{bmatrix} \boldsymbol{u}_i[k] \\ \boldsymbol{u}_i[k+1] \\ \vdots \\ \boldsymbol{u}_i[k+N_h-1] \end{bmatrix}, \tilde{\boldsymbol{w}}_i = \begin{bmatrix} \boldsymbol{w}_i[k] \\ \boldsymbol{w}_i[k+1] \\ \vdots \\ \boldsymbol{w}_i[k+N_h-1] \end{bmatrix}, \tilde{\boldsymbol{v}}_i = \begin{bmatrix} \boldsymbol{v}_i[k] \\ \boldsymbol{v}_i[k+1] \\ \vdots \\ \boldsymbol{v}_i[k+N_h-1] \end{bmatrix},$$

in which k is the current sampling time and N_h is the prediction horizon. Moreover, the compact matrices required for the state-space descriptions and the constraint equations are defined as:

$$\mathbf{A}_i = \begin{bmatrix} A_i \\ A_i^2 \\ \vdots \\ A_i^{N_h} \end{bmatrix}, \ \mathbf{B}_i = \begin{bmatrix} B_i & 0 & 0 & \cdots & 0 \\ A_iB_i & B_i & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ A_i^{N_h-1}B_i & A_i^{N_h-2}B_i & A_i^{N_h-3}B_i & \cdots & B_i \end{bmatrix},$$

$$\mathbf{H}_i = \begin{bmatrix} H_i & 0 & 0 & \cdots & 0 \\ A_iH_i & H_i & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ A_i^{N_h-1}H_i & A_i^{N_h-2}H_i & A_i^{N_h-3}H_i & \cdots & H_i \end{bmatrix},$$

$$\mathbf{\Phi}_i = \begin{bmatrix} \Psi\Phi_i \\ \Psi\Phi_i^2 \\ \vdots \\ \Psi\Phi_i^{N_h} \end{bmatrix}, \ \mathbf{\Gamma}_i = \begin{bmatrix} \Psi\Gamma_i & 0 & 0 & \cdots & 0 \\ \Psi\Phi_i\Gamma_i & \Psi\Gamma_i & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \Psi\Phi_i^{N_h-1}\Gamma_i & \Psi\Phi_i^{N_h-2}\Gamma_i & \Psi\Phi_i^{N_h-3}\Gamma_i & \cdots & \Psi\Gamma_i \end{bmatrix}.$$

B-2 Local Constraints in Chapter 2

This section explains the construction of matrices $\mathbf{E}_i, \mathbf{F}_i, \mathbf{P}_i$ of the first case study (Chapter 2). We collect all the local constraints as a single inequality constraint in the form of $\mathbf{E}_i\tilde{\boldsymbol{u}}_i + \mathbf{F}_i\tilde{\boldsymbol{v}}_i + \mathbf{P}_i \leq 0$.

The startup cost constraints (2-7) can be written as

$$\mathbf{E}_{1,i}\tilde{\boldsymbol{u}}_i + \mathbf{F}_{1,i}\tilde{\boldsymbol{v}}_i + \mathbf{P}_{1,i} \leq 0$$

where

$$\mathbf{E}_{1,i} = \begin{bmatrix} E_{1,i} & 0 & \cdots & 0 \\ 0 & E_{1,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & E_{1,i} \end{bmatrix}, \quad \mathbf{F}_{1,i} = \begin{bmatrix} F_{1,i} & 0 & \cdots & 0 & 0 \\ -F_{1,i} & F_{1,i} & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -F_{1,i} & F_{1,i} \end{bmatrix},$$

$$E_{1,i} = \begin{bmatrix} 0^{2\times 6} & -I & 0^{2\times 2} & 0^{2\times|\mathcal{N}_{-i}|} \\ 0^{2\times 6} & -I & 0^{2\times 2} & 0^{2\times|\mathcal{N}_{-i}|} \end{bmatrix}, \quad F_{1,i} = \begin{bmatrix} \Lambda_i^{\mathbf{su}} \\ 0^{2\times 2} \end{bmatrix},$$

and if $k \geq 1$, $\mathbf{P}_{1,i} = \begin{bmatrix} \boldsymbol{v}_i[k-1] \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ otherwise $\mathbf{P}_{1,i} = 0$.

The capacity constraints (2-8), (2-9), (2-10), (2-11) and (2-12) are written as

$$\mathbf{E}_{2,i}\tilde{\boldsymbol{u}}_i + \mathbf{F}_{2,i}\tilde{\boldsymbol{v}}_i + \mathbf{P}_{2,i} \leq 0$$

where

$$\mathbf{E}_{2,i} = \begin{bmatrix} E_{2,i} & 0 & \cdots & 0 \\ 0 & E_{2,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & E_{2,i} \end{bmatrix}, \quad \mathbf{F}_{2,i} = \begin{bmatrix} F_{2,i} & 0 & \cdots & 0 \\ 0 & F_{2,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & F_{2,i} \end{bmatrix}, \mathbf{P}_{2,i} = \begin{bmatrix} P_{2,i} \\ P_{2,i} \\ \vdots \\ P_{2,i} \end{bmatrix},$$

$$E_{2,i} = \begin{bmatrix} -1 & 0^{1\times 9} & 0^{1\times|\mathcal{N}_{-i}|} \\ 1 & 0^{1\times 9} & 0^{1\times|\mathcal{N}_{-i}|} \\ 0^{1\times 3} & -1 & 0^{1\times(6+|\mathcal{N}_{-i}|)} \\ 0^{1\times 3} & 1 & 0^{1\times(6+|\mathcal{N}_{-i}|)} \\ 0^{1\times 4} & -1 & 0^{1\times(5+|\mathcal{N}_{-i}|)} \\ 0^{1\times 4} & 1 & 0^{1\times(5+|\mathcal{N}_{-i}|)} \\ 0^{1\times 5} & -1 & 0^{1\times(4+|\mathcal{N}_{-i}|)} \\ 0^{1\times 5} & 1 & 0^{1\times(4+|\mathcal{N}_{-i}|)} \\ 0^{|\mathcal{N}_{-i}|\times 10} & & -I \\ 0^{|\mathcal{N}_{-i}|\times 10} & & I \end{bmatrix}, F_{2,i} = \begin{bmatrix} p_{\mathbf{g},i}^{\min} & 0 \\ -p_{\mathbf{g},i}^{\max} & 0 \\ h_{\mathbf{g},i}^{\min} & 0 \\ -h_{\mathbf{g},i}^{\max} & 0 \\ 0 & h_{\mathbf{b},i}^{\min} \\ 0 & -h_{\mathbf{b},i}^{\max} \\ 0^{(2+|\mathcal{N}_{-i}|)\times 1} & 0^{(2+|\mathcal{N}_{-i}|)\times 1} \end{bmatrix},$$

$$P_{2,i} = \begin{bmatrix} 0^{1\times 6} & h_{\mathbf{im},i}^{\min} & -h_{\mathbf{im},i}^{\max} & h_{\mathbf{xc}}^{\min} & \cdots & h_{\mathbf{xc}}^{\min} & -h_{\mathbf{xc}}^{\max} & \cdots & -h_{\mathbf{xc}}^{\max} \end{bmatrix}^T.$$

The up and down spinning electrical power constraints (2-13) and (2-14) are compactly written as

$$\mathbf{E}_{3,i}\tilde{\boldsymbol{u}}_i + \mathbf{F}_{3,i}\tilde{\boldsymbol{v}}_i + \mathbf{P}_{3,i} \leq 0$$

where

$$\mathbf{E}_{3,i} = \begin{bmatrix} E_{3,i} & 0 & \cdots & 0 \\ 0 & E_{3,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & E_{3,i} \end{bmatrix}, \quad \mathbf{F}_{3,i} = 0, \quad \mathbf{P}_{3,i} = \begin{bmatrix} P_{3,i} \\ P_{3,i} \\ \vdots \\ P_{3,i} \end{bmatrix},$$

$$E_{3,i} = \begin{bmatrix} -1 & 0 & 1 & 0^{1\times(7+|\mathcal{N}_{-i}|)} \\ 1 & -1 & 0 & 0^{1\times(7+|\mathcal{N}_{-i}|)} \\ 0 & -1 & 0 & 0^{1\times(7+|\mathcal{N}_{-i}|)} \\ 0 & 0 & -1 & 0^{1\times(7+|\mathcal{N}_{-i}|)} \end{bmatrix}, \quad P_{3,i} = \begin{bmatrix} p_{\mathbf{d},i} \\ -p_{\mathbf{d},i} \\ 0 \\ 0 \end{bmatrix}.$$

The ramping capacity constraints can be expressed as

$$\mathbf{E}_{4,i}\tilde{\boldsymbol{u}}_i + \mathbf{F}_{4,i}\tilde{\boldsymbol{v}}_i + \mathbf{P}_{4,i} \leq 0$$

where

if $k > 1$,

$$\mathbf{E}_{4,i} = \begin{bmatrix} E_{4,i} & 0 & \cdots & 0 & 0 \\ -E_{4,i} & E_{4,i} & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -E_{4,i} & E_{4,i} \end{bmatrix}, \quad \mathbf{F}_{4,i} = 0, \quad \mathbf{P}_{4,i} = \begin{bmatrix} P_{4,i} + P_{4\text{ad},i} \\ P_{4,i} \\ \vdots \\ P_{4,i} \end{bmatrix},$$

otherwise

$$\mathbf{E}_{4,i} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ -E_{4,i} & E_{4,i} & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -E_{4,i} & E_{4,i} \end{bmatrix}, \quad \mathbf{F}_{4,i} = 0, \quad \mathbf{P}_{4,i} = \begin{bmatrix} 0 \\ P_{4,i} \\ \vdots \\ P_{4,i} \end{bmatrix},$$

Furthermore,

$$E_{4,i} = \begin{bmatrix} -1 & 0^{1\times(9+|\mathcal{N}_{-i}|)} \\ 1 & 0^{1\times(9+|\mathcal{N}_{-i}|)} \end{bmatrix}, \quad P_{4,i} = \begin{bmatrix} -p_{\mathbf{g},i}^{\mathbf{down}} \\ -p_{\mathbf{g},i}^{\mathbf{up}} \end{bmatrix}, \quad P_{4\text{ad},i} = \begin{bmatrix} -p_{\mathbf{g},i}[k-1] \\ -p_{\mathbf{g},i}[k-1] \end{bmatrix}.$$

The status change constraints of the production units (the boilers and micro-CHP) are put together as the following. First, inequalities in (2-16) and (2-19) can be written as

$$\mathbf{E}_{5a,i}\tilde{\boldsymbol{u}}_i + \mathbf{F}_{5a,i}\tilde{\boldsymbol{v}}_i + \mathbf{P}_{5a,i} \leq 0$$

where

$$\mathbf{E}_{5a,i} = \begin{bmatrix} E_{5a,i} & 0 & \cdots & 0 \\ 0 & E_{5a,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & E_{5a,i} \end{bmatrix}, \quad \mathbf{F}_{5a,i} = \begin{bmatrix} F_{5a,i} & 0 & \cdots & 0 & 0 \\ -F_{5a,i} & F_{5a,i} & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -F_{5a,i} & F_{5a,i} \end{bmatrix},$$

$$E_{5a,i} = \begin{bmatrix} 0^{2\times 8} & -I & 0^{2\times|\mathcal{N}_{-i}|} \\ 0^{2\times 8} & -I & 0^{2\times|\mathcal{N}_{-i}|} \end{bmatrix}, \quad F_{5a,i} = \begin{bmatrix} I \\ 0^{2\times 2} \end{bmatrix},$$

and if $k \geq 1$, $\mathbf{P}_{5a,i} = \begin{bmatrix} \boldsymbol{v}_i[k-1] \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ otherwise $\mathbf{P}_{5a,i} = 0$.

Equation (2-17) can be written as

$$\mathbf{E}_{5b,i}\tilde{\boldsymbol{u}}_i + \mathbf{F}_{5b,i}\tilde{\boldsymbol{v}}_i + \mathbf{P}_{5b,i} \leq 0$$

where

$$\mathbf{E}_{5b,i} = \begin{bmatrix} E_{5b,i} & 0 & 0 & 0 & \cdots & 0 \\ E_{5b,i} & E_{5b,i} & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \cdots & 0 \\ E_{5b,i} & \cdots & E_{5b,i} & 0 & \cdots & 0 \\ 0 & E_{5b,i} & \cdots & E_{5b,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & E_{5b,i} & \cdots & E_{5b,i} \end{bmatrix}, \quad \mathbf{P}_{5b,i} = \begin{bmatrix} \sum_{l=k+1-\Delta k_{\mathbf{up,g}}}^{k-1} z_{\mathbf{g},i}[l] \\ \sum_{l=k+2-\Delta k_{\mathbf{up,g}}}^{k-1} z_{\mathbf{g},i}[l] \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$\mathbf{F}_{5b,i} = \begin{bmatrix} F_{5b,i} & 0 & \cdots & 0 \\ 0 & F_{5b,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & F_{5b,i} \end{bmatrix}, \quad \forall k \geq \Delta k_{\mathbf{up,g}},$$

in which

$$E_{5b,i} = \begin{bmatrix} 0^{1\times 8} & 1 & 0 & 0^{2\times|\mathcal{N}_{-i}|} \end{bmatrix}, \quad F_{5b,i} = \begin{bmatrix} -1 & 0 \end{bmatrix}.$$

Similarly, Equation (2-20) which refers to the boilers can be written as

$$\mathbf{E}_{5c,i}\tilde{\boldsymbol{u}}_i + \mathbf{F}_{5c,i}\tilde{\boldsymbol{v}}_i + \mathbf{P}_{5c,i} \leq 0$$

where

$$\mathbf{E}_{5c,i} = \begin{bmatrix} E_{5c,i} & 0 & 0 & 0 & \cdots & 0 \\ E_{5c,i} & E_{5c,i} & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \cdots & 0 \\ E_{5c,i} & \cdots & E_{5c,i} & 0 & \cdots & 0 \\ 0 & E_{5c,i} & \cdots & E_{5c,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & E_{5c,i} & \cdots & E_{5c,i} \end{bmatrix}, \quad \mathbf{P}_{5c,i} = \begin{bmatrix} \sum_{l=k+1-\Delta k_{\mathbf{up,b}}}^{k-1} z_{\mathbf{b},i}[l] \\ \sum_{l=k+2-\Delta k_{\mathbf{up,b}}}^{k-1} z_{\mathbf{b},i}[l] \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$\mathbf{F}_{5c,i} = \begin{bmatrix} F_{5c,i} & 0 & \cdots & 0 \\ 0 & F_{5c,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & F_{5c,i} \end{bmatrix}, \quad \forall k \geq \Delta k_{\mathbf{up,b}},$$

in which

$$E_{5c,i} = \begin{bmatrix} 0^{1\times 8} & 0 & 1 & 0^{2\times|\mathcal{N}_{-i}|} \end{bmatrix}, \quad F_{5c,i} = \begin{bmatrix} 0 & -1 \end{bmatrix}.$$

Meanwhile, Equation (2-18) can be expressed as

$$\mathbf{E}_{5d,i}\tilde{\boldsymbol{u}}_i + \mathbf{F}_{5d,i}\tilde{\boldsymbol{v}}_i + \mathbf{P}_{5d,i} \leq 0$$

where

$$\mathbf{E}_{5d,i} = \begin{bmatrix} E_{5d,i} & \cdots & E_{5d,i} & 0 & 0 & \cdots & 0 \\ 0 & E_{5d,i} & \cdots & E_{5d,i} & 0 & \cdots & 0 \\ 0 & 0 & E_{5d,i} & \cdots & E_{5d,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & E_{5d,i} & \cdots & E_{5d,i} \end{bmatrix}, \quad \mathbf{P}_{5d,i} = \begin{bmatrix} v_{\mathbf{g},i}[k-1]-1 \\ -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix},$$

$$\mathbf{F}_{5d,i} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ F_{5d,i} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & F_{5d,i} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & F_{5c,i} & 0 & \cdots & 0 \end{bmatrix}, \quad \mathbf{P}_{5d,i} \in \mathbb{R}^{(N_h - \Delta k_{\mathbf{down,g},i})}, \forall k > 1,$$

in which,

$$E_{5d,i} = E_{5b,i}, \quad F_{5b,i} = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

Similarly, Equation (2-21) can be expressed as

$$\mathbf{E}_{5e,i}\tilde{\boldsymbol{u}}_i + \mathbf{F}_{5e,i}\tilde{\boldsymbol{v}}_i + \mathbf{P}_{5e,i} \leq 0$$

where

$$\mathbf{E}_{5e,i} = \begin{bmatrix} E_{5e,i} & \cdots & E_{5e,i} & 0 & 0 & \cdots & 0 \\ 0 & E_{5e,i} & \cdots & E_{5e,i} & 0 & \cdots & 0 \\ 0 & 0 & E_{5e,i} & \cdots & E_{5e,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & E_{5e,i} & \cdots & E_{5e,i} \end{bmatrix}, \quad \mathbf{P}_{5e,i} = \begin{bmatrix} v_{\mathbf{b},i}[k-1]-1 \\ -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix},$$

$$\mathbf{F}_{5e,i} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ F_{5e,i} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & F_{5e,i} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & F_{5c,i} & 0 & \cdots & 0 \end{bmatrix}, \quad \mathbf{P}_{5e,i} \in \mathbb{R}^{(N_h - \Delta k_{\mathbf{down},\mathbf{g},i})}, \forall k > 1,$$

in which,

$$E_{5e,i} = E_{5c,i}, \quad F_{5b,i} = \begin{bmatrix} 0 & 1 \end{bmatrix}.$$

Therefore we can obtain

$$\mathbf{E}_{5,i} = \begin{bmatrix} \mathbf{E}_{5a,i}^T & \mathbf{E}_{5b,i}^T & \mathbf{E}_{5c,i}^T & \mathbf{E}_{5d,i}^T & \mathbf{E}_{5e,i}^T \end{bmatrix}^T,$$
$$\mathbf{F}_{5,i} = \begin{bmatrix} \mathbf{F}_{5a,i}^T & \mathbf{F}_{5b,i}^T & \mathbf{F}_{5c,i}^T & \mathbf{F}_{5d,i}^T & \mathbf{F}_{5e,i}^T \end{bmatrix}^T,$$
$$\mathbf{P}_{5,i} = \begin{bmatrix} \mathbf{P}_{5a,i}^T & \mathbf{P}_{5b,i}^T & \mathbf{P}_{5c,i}^T & \mathbf{P}_{5d,i}^T & \mathbf{P}_{5e,i}^T \end{bmatrix}^T.$$

Finally, we obtain the matrices for the local constraints as

$$\mathbf{E} = \begin{bmatrix} \mathbf{E}_{1,i}^T & \mathbf{E}_{2,i}^T & \mathbf{E}_{3,i}^T & \mathbf{E}_{4,i}^T & \mathbf{E}_{5,i}^T \end{bmatrix}^T,$$
$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{1,i}^T & \mathbf{F}_{2,i}^T & \mathbf{F}_{3,i}^T & \mathbf{F}_{4,i}^T & \mathbf{F}_{5,i}^T \end{bmatrix}^T,$$
$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{1,i}^T & \mathbf{P}_{2,i}^T & \mathbf{P}_{3,i}^T & \mathbf{P}_{4,i}^T & \mathbf{P}_{5,i}^T \end{bmatrix}^T.$$

B-3 Balance Constraints in Chapter 2

Here we provide the example of constructing matrices G_{ij} and K_i which appear in the compact forms of the balance constraint. Consider a grid with 4 agents as depicted in Figure B-1. The balance constraints of this system can be written as:

$$h_{\mathbf{xc},12} + h_{\mathbf{xc},21} = 0,$$
$$h_{\mathbf{xc},31} + h_{\mathbf{xc},13} = 0,$$
$$h_{\mathbf{xc},23} + h_{\mathbf{xc},32} = 0,$$
$$h_{\mathbf{xc},24} + h_{\mathbf{xc},42} = 0.$$

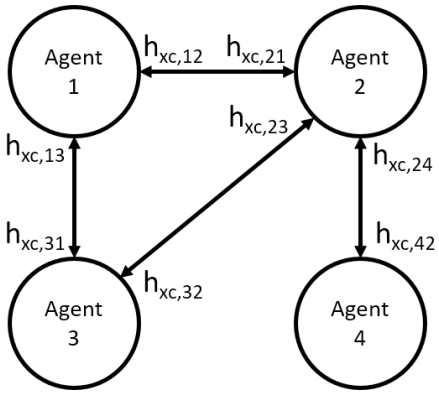


Figure B-1: The example of a 4-agent grid system with the possibility of exchanging energy.

By defining

$$\boldsymbol{u}_{c,1} = \begin{bmatrix} h_{\mathbf{xc},12} & h_{\mathbf{xc},13} \end{bmatrix}^T, \boldsymbol{u}_{c,2} = \begin{bmatrix} h_{\mathbf{xc},21} & h_{\mathbf{xc},23} & h_{\mathbf{xc},24} \end{bmatrix}^T,$$
$$\boldsymbol{u}_{c,3} = \begin{bmatrix} h_{\mathbf{xc},31} & h_{\mathbf{xc},32} \end{bmatrix}^T, \boldsymbol{u}_{c,4} = h_{\mathbf{xc},42},$$

the matrices $G_{ij}\ \forall j \in \mathcal{N}_{-i}$ which correspond to those balance constraints are

$$G_{12} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ G_{13} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$
$$G_{21} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},\ G_{23} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},\ G_{24} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$
$$G_{13} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ G_{32} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$
$$G_{42} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}.$$

Whereas, for the compact format of the formulation of the ADMM method which is based on the exchange problem, the matrices $K_i\ \forall i \in \mathcal{N}$ are defined as

$$K_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},\ K_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ K_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},\ K_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

B-4 Local Constraints in Chapter 3

This section explains the construction of matrices $\mathbf{E}_i, \mathbf{F}_i, \mathbf{P}_i$ of the second case study (Chapter 3). They are formed from the startup cost constraints (3-23) and the capacity constraints (3-24), (3-25), (3-26), (3-27), (3-28) and (3-29).

The startup cost constraints (3-23) can be written as

$$\mathbf{E}_{1,i}\tilde{\boldsymbol{u}}_i + \mathbf{F}_{1,i}\tilde{\boldsymbol{v}}_i + \mathbf{P}_{1,i} \leq 0$$

where

$$\mathbf{E}_{1,i} = \begin{bmatrix} E_{1,i} & 0 & \cdots & 0 \\ 0 & E_{1,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & E_{1,i} \end{bmatrix}, \quad \mathbf{F}_{1,i} = \begin{bmatrix} F_{1,i} & 0 & \cdots & 0 & 0 \\ -F_{1,i} & F_{1,i} & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -F_{1,i} & F_{1,i} \end{bmatrix},$$

$$E_{1,i} = \begin{bmatrix} 0^{2\times 6} & -I \\ 0^{2\times 6} & -I \end{bmatrix}, \quad F_{1,i} = \begin{bmatrix} \Lambda_i^{\mathbf{su}} \\ 0^{2\times 2} \end{bmatrix},$$

and if $k \geq 1$, $\mathbf{P}_{1,i} = \begin{bmatrix} \boldsymbol{v}_i[k-1] \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ otherwise $\mathbf{P}_{1,i} = 0$.

Meanwhile, the capacity constraints can be written as

$$\mathbf{E}_{2,i}\tilde{\boldsymbol{u}}_i + \mathbf{F}_{2,i}\tilde{\boldsymbol{v}}_i + \mathbf{P}_{2,i} \leq 0$$

where

$$\mathbf{E}_{2,i} = \begin{bmatrix} E_{2,i} & 0 & \cdots & 0 \\ 0 & E_{2,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & E_{2,i} \end{bmatrix}, \quad \mathbf{F}_{2,i} = \begin{bmatrix} F_{2,i} & 0 & \cdots & 0 \\ 0 & F_{2,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & F_{2,i} \end{bmatrix}, \mathbf{P}_{2,i} = \begin{bmatrix} P_{2,i} \\ P_{2,i} \\ \vdots \\ P_{2,i} \end{bmatrix},$$

$$E_{2,i} = \begin{bmatrix} -1 & 0 & 0^{1\times 2} & 0^{1\times 2} & 0^{1\times 2} \\ 1 & 0 & 0^{1\times 2} & 0^{1\times 2} & 0^{1\times 2} \\ 0 & -1 & 0^{1\times 2} & 0^{1\times 2} & 0^{1\times 2} \\ 0 & 1 & 0^{1\times 2} & 0^{1\times 2} & 0^{1\times 2} \\ 0^{1\times 2} & -1 & 0 & 0^{1\times 2} & 0^{1\times 2} \\ 0^{1\times 2} & 1 & 0 & 0^{1\times 2} & 0^{1\times 2} \\ 0^{1\times 2} & 0 & -1 & 0^{1\times 2} & 0^{1\times 2} \\ 0^{1\times 2} & 0 & 1 & 0^{1\times 2} & 0^{1\times 2} \\ 0^{1\times 2} & 0^{1\times 2} & -1 & 0 & 0^{1\times 2} \\ 0^{1\times 2} & 0^{1\times 2} & 1 & 0 & 0^{1\times 2} \\ 0^{1\times 2} & 0^{1\times 2} & 0 & -1 & 0^{1\times 2} \\ 0^{1\times 2} & 0^{1\times 2} & 0 & 1 & 0^{1\times 2} \end{bmatrix}, F_{2,i} = \begin{bmatrix} h_{\mathbf{b},i}^{\min} & 0 \\ -h_{\mathbf{b},i}^{\max} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & c_{\mathbf{ch},i}^{\min} \\ 0 & -c_{\mathbf{ch},i}^{\max} \\ 0^{6\times 1} & 0^{6\times 1} \end{bmatrix},$$

$$P_{2,i} = \begin{bmatrix} 0^{1\times 2} & 0 & -h_{\mathbf{im},i}^{\max} & 0^{1\times 2} & 0 & -c_{\mathbf{im},i}^{\max} & 0 & -u_{\mathbf{sh},i}^{\max} & 0 & -u_{\mathbf{sc},i}^{\max} \end{bmatrix}^T.$$

Hence, we obtain the matrices for the local constraints as

$$\mathbf{E} = \begin{bmatrix} \mathbf{E}_{1,i}^T & \mathbf{E}_{2,i}^T \end{bmatrix}^T,$$
$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{1,i}^T & \mathbf{F}_{2,i}^T \end{bmatrix}^T,$$
$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{1,i}^T & \mathbf{P}_{2,i}^T \end{bmatrix}^T.$$

B-5 Coupling Constraints in Chapter 3

The coupling constraints in the numerical study of Chapter 3 are written as

$$\begin{aligned}
&V_{\mathbf{h},1}(t) + V_{\mathbf{c},2}(t) \leq \bar{V}_{\mathbf{hc},12},\\
&V_{\mathbf{h},2}(t) + V_{\mathbf{c},3}(t) \leq \bar{V}_{\mathbf{hc},23},\\
&V_{\mathbf{h},3}(t) + V_{\mathbf{c},1}(t) \leq \bar{V}_{\mathbf{hc},31}, \forall t \in \{k+1, \cdots, N_h\}.
\end{aligned}$$

This constraints can be written compactly as in (3-35f) by defining

$$\begin{aligned}
G_{11} = G_{22} = G_{33} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},\\
G_{12} = G_{23} = G_{31} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},\\
G_{13} = G_{21} = G_{32} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}$$

Appendix C

Distributed Optimization Based on Dual Decomposition

This appendix discusses some distributed optimization methods which are mainly based on the Lagrange dual theory. Section C-1 provides some important theories in convex optimization which are the bases whereas the methods are discussed in Section C-2.

C-1 Preliminaries

In this section, a summary of convex optimization and Lagrange duality that are important and related to the thesis are presented. They are extracted from [44] and [45].

C-1-1 Convex Optimization

Local and Global Optima of a Function

Consider the following general optimization problem

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize or maximize}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{x} \in \mathcal{P}, \end{aligned} \tag{C-1}$$

and the ball centered at $\bar{\boldsymbol{x}}$ with radius ϵ which is the set

$$B(\bar{\boldsymbol{x}}, \epsilon) := \{\boldsymbol{x} | ||\boldsymbol{x} - \bar{\boldsymbol{x}}|| \leq \epsilon\}. \tag{C-2}$$

We have the following definitions of local/global, strict/non-strict minima/maxima:

Definition C-1.1. *$\boldsymbol{x} \in \mathcal{P}$ is a local minimum of C-1 if there exists $\epsilon > 0$ such that $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ for all $\boldsymbol{y} \in B(\boldsymbol{x}, \epsilon) \cap \mathcal{P}$.*

Definition C-1.2. *$\boldsymbol{x} \in \mathcal{P}$ is a global minimum of C-1 if $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ for all $\boldsymbol{y} \in \mathcal{P}$.*

Definition C-1.3. *$\boldsymbol{x} \in \mathcal{P}$ is a strict local minimum of C-1 if there exists $\epsilon > 0$ such that $f(\boldsymbol{x}) < f(\boldsymbol{y})$ for all $\boldsymbol{y} \in B(\boldsymbol{x}, \epsilon) \cap \mathcal{P},\ \boldsymbol{y} \neq \boldsymbol{x}$.*

Definition C-1.4. *$\boldsymbol{x} \in \mathcal{P}$ is a strict global minimum of C-1 if $f(\boldsymbol{x}) < f(\boldsymbol{y})$ for all $\boldsymbol{y} \in \mathcal{P},\ \boldsymbol{y} \neq \boldsymbol{x}$.*

Definition C-1.5. *$\boldsymbol{x} \in \mathcal{P}$ is a local maximum of C-1 if there exists $\epsilon > 0$ such that $f(\boldsymbol{x}) \geq f(\boldsymbol{y})$ for all $\boldsymbol{y} \in B(\boldsymbol{x}, \epsilon) \cap \mathcal{P}$.*

Definition C-1.6. *$\boldsymbol{x} \in \mathcal{P}$ is a global maximum of C-1 if $f(\boldsymbol{x}) \geq f(\boldsymbol{y})$ for all $\boldsymbol{y} \in \mathcal{P}$.*

Definition C-1.7. *$\boldsymbol{x} \in \mathcal{P}$ is a strict local maximum of C-1 if there exists $\epsilon > 0$ such that $f(\boldsymbol{x}) > f(\boldsymbol{y})$ for all $\boldsymbol{y} \in B(\boldsymbol{x}, \epsilon) \cap \mathcal{P},\ \boldsymbol{y} \neq \boldsymbol{x}$.*

Definition C-1.8. *$\boldsymbol{x} \in \mathcal{P}$ is a strict global maximum of C-1 if $f(\boldsymbol{x}) > f(\boldsymbol{y})$ for all $\boldsymbol{y} \in \mathcal{P},\ \boldsymbol{y} \neq \boldsymbol{x}$.*

Convex Sets and Functions

The following is the definition of a convex set:

Definition C-1.9. *A subset $\mathcal{C} \in \mathbb{R}^n$ is a convex set if the line segment between any two points in $\mathcal{C}$ lies in $\mathcal{C}$, that is*

$$\theta \boldsymbol{x} + (1 - \theta)\boldsymbol{y} \in \mathcal{C},$$

for any $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{C}$ and $0 \leq \theta \leq 1$.

This leads to the following propositions:

Proposition C-1.1. *If $\mathcal{S}, \mathcal{T}$ are convex sets, then $\mathcal{S} \cap \mathcal{T}$ is a convex set.*

Proposition C-1.2. *The intersection of any collection of convex sets is a convex set.*

Convexity of a function is defined as follows

Definition C-1.10 (Convexity)**.** *A function $f : \mathbb{R}^n \to \mathbb{R}$ is* convex *if*

$$f(\theta \boldsymbol{x} + (1 - \theta)\boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + (1 - \theta) f(\boldsymbol{y}), \tag{C-3}$$

for any $\boldsymbol{x}, \boldsymbol{y}$ and θ with $0 \leq \theta \leq 1$.

Definition C-1.11 (Strict Convexity)**.** *A function $f : \mathbb{R}^n \to \mathbb{R}$ is* strictly convex *if*

$$f(\theta \boldsymbol{x} + (1 - \theta)\boldsymbol{y}) < \theta f(\boldsymbol{x}) + (1 - \theta) f(\boldsymbol{y}), \tag{C-4}$$

for any $\boldsymbol{x}, \boldsymbol{y}$, $\boldsymbol{x} \neq \boldsymbol{y}$ and θ with $0 < \theta < 1$.

Now consider the following optimization problem, where the feasible region is simply described as the set $\mathcal{P}$:

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{x} \in \mathcal{P}, \end{aligned} \tag{C-5}$$

Theorem C-1.1. *Suppose that $\mathcal{P}$ is a convex set, $f : \mathcal{P} \to \mathbb{R}$ is a convex function, and $\bar{\boldsymbol{x}}$ is a local minimum of* (C-5)*. Then $\bar{\boldsymbol{x}}$ is a global minimum of f over $\mathcal{P}$.*

Proof. Suppose $\bar{\boldsymbol{x}}$ is not a global minimum, i.e., there exists $\boldsymbol{y} \in \mathcal{P}$ for which $f(\boldsymbol{y}) < f(\bar{\boldsymbol{x}})$. Let $\boldsymbol{y}(\theta) = \theta\bar{\boldsymbol{x}} + (1-\theta)\boldsymbol{y}$, which is a convex combination of $\bar{\boldsymbol{x}}$ and $\boldsymbol{y}$ for $\theta \in [0,1]$ and therefore $\boldsymbol{y}(\theta) \in \mathcal{P}$ for $\theta \in [0,1]$. Note that $\boldsymbol{y}(\theta) \to \bar{\boldsymbol{x}}$ as $\theta \to 1$. From the convexity of $f(\cdot)$,

$$f(\boldsymbol{y}(\theta)) = f(\theta\bar{\boldsymbol{x}} + (1+\theta)\boldsymbol{y}) \leq \theta f(\bar{\boldsymbol{x}}) + (1+\theta)f(\boldsymbol{y}) < \theta f(\bar{\boldsymbol{x}}) + (1+\theta)f(\bar{\boldsymbol{x}}) = f(\bar{\boldsymbol{x}}),$$

for all $\theta \in (0,1)$. Therefore $f(\boldsymbol{y}(\theta)) < f(\bar{\boldsymbol{x}})$ for all $\theta \in (0,1)$ and so $\bar{\boldsymbol{x}}$ is not a local minimum, resulting in a contradiction. □

Concave Functions and Maximization

A concave function is defined as the following.

Definition C-1.12. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is* concave *if*

$$f(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) \geq \theta f(\boldsymbol{x}) + (1-\theta)f(\boldsymbol{y}), \tag{C-6}$$

for any $\boldsymbol{x}, \boldsymbol{y}$ and θ with $0 \leq \theta \leq 1$.

Definition C-1.13. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is* strictly concave *if*

$$f(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) > \theta f(\boldsymbol{x}) + (1-\theta)f(\boldsymbol{y}), \tag{C-7}$$

for any $\boldsymbol{x}, \boldsymbol{y}$, $\boldsymbol{x} \neq \boldsymbol{y}$ and θ with $0 < \theta < 1$.

Now consider the following optimization problem, where the feasible region is simply described as the set $\mathcal{P}$:

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{maximize}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{x} \in \mathcal{P}, \end{aligned} \tag{C-8}$$

Theorem C-1.2. *Suppose that $\mathcal{P}$ is a convex set, $f : \mathcal{P} \to \mathbb{R}$ is a concave function, and $\bar{\boldsymbol{x}}$ is a local maximum of* (C-8)*. Then $\bar{\boldsymbol{x}}$ is a global maximum of f over $\mathcal{P}$.*

Affine functions

An important class of functions which is related to our thesis is a class of affine functions. An affine function is defined as $f(\boldsymbol{x}) = a^T\boldsymbol{x} + b$.

Proposition C-1.3. *An affine function is both convex and concave.*

Proposition C-1.4. *if $f(\boldsymbol{x})$ is both convex and concave, then $f(\boldsymbol{x})$ is an affine function.*

Convex Programs

Consider the following optimization problem which can be either a minimization or maximization problem:

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize or maximize}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{x} \in \mathcal{P}, \end{aligned} \tag{C-9}$$

Definition C-1.14. *We call Problem* (C-9) *a* convex program *if*

1. *$\mathcal{P}$ is a convex set, and*
2. *we are minimizing and $f(\boldsymbol{x})$ is a convex function or we are maximizing and $f(\boldsymbol{x})$ is a concave function.*

Therefore, from the above theorems, every local optimia of the objective function of a convex program is a global optima of the objective function.

A Separating Hyperplane

The separating hyperplane theorem for convex sets is the main tool which is used in developing duality and analyzing dual problem. The theorem is stated as follows:

Theorem C-1.3 (Strong Separating Hyperplane Theorem)**.** *Suppose that $\mathcal{S}$ is a convex set in $\mathbb{R}^n$ and there is a point $\bar{\boldsymbol{x}} \notin \mathcal{S}$. Then there exists a vector $u \neq 0$ and a scalar α for which the following hold:*

1. *$u^T\boldsymbol{x} > \alpha \ \forall \boldsymbol{x} \in \mathcal{S}$.*
2. *$u^T\boldsymbol{x} < \alpha \ \forall \boldsymbol{x} \in \mathcal{S}$.*

A weaker version of this theorem is as follows:

Theorem C-1.4 (Weak Separating Hyperplane Theorem)**.** *Suppose that $\mathcal{S}$ is a convex set in $\mathbb{R}^n$ and there is a point $\bar{\boldsymbol{x}} \notin \mathcal{S}$ or $\bar{\boldsymbol{x}} \in \partial\mathcal{S}$. Then there exists a vector $u \neq 0$ and a scalar α for which the following hold:*

1. *$u^T\boldsymbol{x} \geq \alpha \ \forall \boldsymbol{x} \in \mathcal{S}$.*
2. *$u^T\boldsymbol{x} \leq \alpha \ \forall \boldsymbol{x} \in \mathcal{S}$.*

C-1-2 Lagrange Duality Theory

Consider the following optimization problem

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad J(\boldsymbol{x}) \tag{C-10a}$$

$$\text{subject to} \quad g_i(\boldsymbol{x}) \leq 0, \quad i = 1, ..., N_q, \tag{C-10b}$$

$$h_i(\boldsymbol{x}) = 0, \quad i = 1, ..., N_r, \tag{C-10c}$$

$$\boldsymbol{x} \in \mathcal{P}, \tag{C-10d}$$

where $\boldsymbol{x} \in \mathbb{R}^n$, $J : \mathbb{R}^n \to \mathbb{R}$ is the cost function, Equation (C-10b) is the inequality constraints in which $g_i : \mathbb{R}^n \to \mathbb{R},\ \forall i \in \{1, ..., N_q\}$ are the inequality constraint functions, and Equation (C-10b) is the equality constraint in which $h_i : \mathbb{R}^n \to \mathbb{R},\ \forall i \in \{1, ..., N_r\}$ are the equality constraint functions. Moreover, Equation (C-10d) is an additional constraint where $\mathcal{P}$ is the feasible set in which $\boldsymbol{x}$ must lie. Note it is a general optimization problem. For instance, a maximization problem can also be converted into this form.

Lagrange Dual Function

The Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^{N_q} \times \mathbb{R}^{N_r} \to \mathbb{R}$ associated with the optimization problem (C-10) is defined as

$$L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = J(\boldsymbol{x}) + \sum_{i=1}^{N_q} \lambda_i g_i(\boldsymbol{x}) + \sum_{i=1}^{N_r} \nu_i h_i(\boldsymbol{x}), \tag{C-11}$$

where λ_i is referred as the *Lagrange multiplier* associated with the ith inequality constraint and similarly ν_i is referred as the *Lagrange multiplier* associated with the ith equality constraint. The vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ which consist of λ_i and ν_i $\forall i\{1, 2, \cdots, N\}$ are called the *dual variables* or *Lagrange multiplier vectors* associated with the problem (C-10).

The Lagrange dual function g_d: $\mathbb{R}^{N_q} \times \mathbb{R}^{N_r} \to \mathbb{R}$ is then defined as the minimum value of the Lagrangian L over $\boldsymbol{x}$ for $\boldsymbol{\lambda} \in \mathbb{R}^{N_q}, \boldsymbol{\nu} \in \mathbb{R}^{N_r}$,

$$g_d(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\boldsymbol{x} \in \mathcal{P}} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\boldsymbol{x} \in \mathcal{P}} \left(J(\boldsymbol{x}) + \sum_{i=1}^{N_q} \lambda_i g_i(\boldsymbol{x}) + \sum_{i=1}^{N_r} \nu_i h_i(\boldsymbol{x}) \right). \tag{C-12}$$

Theorem C-1.5. *The dual function $g_d(\boldsymbol{\lambda}, \boldsymbol{\nu})$ is a concave function.*

Proof. The dual function is the pointwise infimum of a family of affine functions of $(\boldsymbol{\lambda}, \boldsymbol{\nu})$, thus it is concave (even if problem (C-10) is not convex). □

Lagrange Dual Problem

The dual problem that is associated with the the problem (C-10) si defined as the following:

$$\begin{aligned} &\underset{\boldsymbol{\lambda}, \boldsymbol{\nu}}{\text{maximize}} && g_d(\boldsymbol{\lambda}, \boldsymbol{\nu}) \\ &\text{subject to} && \boldsymbol{\lambda} \geq 0. \end{aligned} \tag{C-13}$$

The Lagrange dual problem (C-13) is a convex optimization problem since the objective to be maximized is concave and the constraint is convex. This holds whether or not the primal problem is convex [45]. It is an optimization problem in which the objective is to find the tightest lower bound of Problem (C-10) because the dual function gives a lower bound as stated in the following theorem.

Theorem C-1.6 (Weak Duality Theorem)**.** *If $\tilde{\boldsymbol{x}}$ is feasible for Problem* (C-10) *and $\tilde{\boldsymbol{\lambda}}, \tilde{\nu}$ is feasible for Problem* (C-13)*, then:*

$$g_d(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) \leq J(\tilde{\boldsymbol{x}}). \tag{C-14}$$

In particular, let $p^\star$ and $d^\star$ be the optimal solution of the primal problem (C-10) *and the dual problem* (C-13)*, we have*

$$d^\star \leq p^\star. \tag{C-15}$$

Proof. For any feasible $\tilde{\boldsymbol{x}} \in \mathcal{P}$, i.e., that satisfies the inequality and equality constraints in (C-10) and for any pair $(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ with $\tilde{\boldsymbol{\lambda}} \geq 0$, we have

$$\sum_{i=1}^{N_q} \tilde{\boldsymbol{\lambda}}_i g_i(\tilde{\boldsymbol{x}}) + \sum_{i=1}^{N_r} \tilde{\boldsymbol{\nu}}_i h_i(\tilde{\boldsymbol{x}}) \leq 0. \tag{C-16}$$

This implies that

$$g_d(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) = \inf_{\boldsymbol{x} \in \mathcal{P}} L(\boldsymbol{x}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) \leq L(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) \leq J(\tilde{\boldsymbol{x}}), \tag{C-17}$$

where the last inequality follows from Equation (C-12) and Equation (C-17). Therefore $d^\star \leq p^\star$. □

Strong Duality

From the weak duality theorem, we know that $d^\star$ is the best lower bound on $p^\star$. The difference $p^\star - d^\star$ is then called as the *optimal duality gap* of the original problem. *Strong duality* holds if the equality

$$d^\star = p^\star \tag{C-18}$$

holds, which also means the optimal duality gap is zero. The strong duality property infers that the problem can be solved by solving its dual problem. If the primal problem (C-10) is convex, usually the problem have strong duality [45].

Optimality Conditions

Consider the primal optimization problem (C-10) and its dual (C-13). Assume that strong duality holds and the cost functions and constraint functions are differentiable. $x^\star$ and $(\lambda^\star, \nu^\star)$ can be the primal and dual optimal points if they satisfy the following condition:

$$\nabla J(x^\star) + \sum_{i=1}^{N_q} \lambda_i^\star \nabla g_i(x^\star) + \sum_{i=1}^{N_r} \nu_i^\star \nabla h_i(x^\star) = 0, \tag{C-19}$$

$$\lambda_i^\star g_i(x^\star) = 0, \quad i = 1, \ldots, N_q, \tag{C-20}$$

$$\lambda_i^\star \geq 0, \quad i = 1, \ldots, N_q, \tag{C-21}$$

$$g_i(x^\star) \leq 0, \quad i = 1, \ldots, N_q, \tag{C-22}$$

$$h_i(x^\star) = 0, \quad i = 1, \ldots, N_r. \tag{C-23}$$

The KKT conditions are necessary conditions for any primal-dual optimal pair if strong duality holds and the cost and constraints are differentiable. If the primal problem is also convex then the KKT conditions are sufficient. The proof of KKT conditions can be found in [45].

C-2 Dual Decomposition

The main idea of distributed approach is that large-scale problems are decomposable. So in this approach, the problem is decomposed into many smaller subproblems (agents) such that there is a local optimization problem at each agent which is expected to be computationally easier and there exist some level of communication or coordination among the agents in order to improve the overall performance [8], [20], [46]. Based on the Lagrange dual theory, we know that it is possible to solve a problem by solving the associated dual problem. Therefore, there are decomposition methods which are based on the dual theory. These methods are called dual decomposition. In this section, we discuss the method and the corresponding algorithm to solve the problem. Moreover, we also provide some examples and compare it with a different distributed optimization strategy. In the end, we discuss a specific issue arisen when using this method which is related to our main problem and explain an extension of this method which is able to deal with this issue.

C-2-1 Dual Ascent Method

The dual ascent method is an iterative-based optimization algorithm which is developed based on the dual theory. In this method, first we decompose the problem which we want to solve into smaller subproblems and then apply a gradient ascent method to solve the dual problem.

In order to describe the method, first we consider the following optimization problem:

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & \sum_{i=1}^{N} J_i(\boldsymbol{x}_i) && \text{(C-24a)} \\ \text{subject to} \quad & \boldsymbol{x}_i \in \mathcal{X}_i, \quad i = 1, ..., N, && \text{(C-24b)} \\ & \sum_{i=1}^{N} A_i \boldsymbol{x}_i \leq c, && \text{(C-24c)} \end{aligned}$$

where agent i has local variable $\boldsymbol{x}_i \in \mathbb{R}^{n_i}$, local constraint set $\mathcal{X}_i \subseteq \mathbb{R}^{n_i}$, and local cost $J_i(\boldsymbol{x}_i)$. Moreover, the system has a coupling inequality constraint $\sum_{i=1}^{N} A_i \boldsymbol{x}_i \leq c$. In addition, the number of agent is N, and the total cost function is $\sum_{i=1}^{N} J_i(\boldsymbol{x}_i)$. Moreover, it is assumed that the cost function $J_i(\boldsymbol{x}_i)$ for all i are strictly convex.

Due to the existence of the coupling constraint (Equation (C-24c)), this problem is not trivially separable. However, by introducing the Lagrange multiplier $\boldsymbol{\lambda}$ which is associated with the coupling constraint, we can obtain the dual problem which is separable as follows:

$$\begin{aligned} \underset{\boldsymbol{\lambda}}{\text{maximize}} \quad & \underset{\{\boldsymbol{x}_i \in \mathcal{X}_i\}_{i=1}^{N}}{\text{minimize}} \quad \sum_{i=1}^{N} \Big(J_i(\boldsymbol{x}_i) + \boldsymbol{\lambda}^T A_i \boldsymbol{x}_i \Big) - \boldsymbol{\lambda}^T c \\ \text{subject to} \quad & \boldsymbol{\lambda} \geq 0. \end{aligned} \tag{C-25}$$

The dual problem can be decomposed into N minimization subproblems since $\boldsymbol{\lambda}^T c$ is a constant and can be omitted. Note that for simplicity we only deal with one coupling constraint. Obviously we can extend the method for more than one coupling constraint by simply introducing the associated Lagrange multipliers and form the corresponding Lagrangian.

Thus, the dual problem can be solved in two stages. The first stage is solving N subproblems. Each subproblem is written as follows

$$\begin{aligned} \underset{\boldsymbol{x}_i}{\text{minimize}} \quad & \left(J_i(\boldsymbol{x}_i) + \boldsymbol{\lambda}^T A_i \boldsymbol{x}_i\right) \\ \text{subject to} \quad & \boldsymbol{x}_i \in \mathcal{X}_i, \end{aligned} \tag{C-26}$$

with the optimal value of $\boldsymbol{x}_i$ of this problem at each iteration is denoted by $\boldsymbol{x}_i^\star$. The second stage is updating the Lagrange multiplier $\boldsymbol{\lambda}$ as a maximization problem as follows

$$\begin{aligned} \underset{\boldsymbol{\lambda}}{\text{maximize}} \quad & \sum_{i=1}^{N} \left(J_i(\boldsymbol{x}_i^\star) + \boldsymbol{\lambda}^T A_i \boldsymbol{x}_i^\star\right) - \boldsymbol{\lambda}^T c \\ \text{subject to} \quad & \boldsymbol{\lambda} \geq 0. \end{aligned} \tag{C-27}$$

This is a convex problem which can be solved using an ascent method. In this method, the Lagrange multipliers are iteratively updated as follows

$$\boldsymbol{\lambda}^{(q+1)} = \max(0, \boldsymbol{\lambda}^{(q)} + \alpha_g d), \tag{C-28}$$

where (q) denotes the number of iterations, α_g is the step size and $0 \leq \alpha_g \leq 1$ and d is the direction as defined as

$$d = \sum_{i=1}^{N} A_i \boldsymbol{x}_i^\star - c. \tag{C-29}$$

The stopping criterion of this method is taken from one of the optimality conditions, that is complementary slackness (Equation (C-20)). So the iteration stops if $\boldsymbol{\lambda}^T(\sum_{i=1}^{N} A_i \boldsymbol{x}_i^\star - c) \leq \epsilon$ for a small value of $\epsilon > 0$. The dual ascent method for this problem is summarized in Algorithm 5.

The dual decomposition methods will find global optimal if strong duality holds. Therefore convexity of the problem is important. This can be see in [22], [23], [24] in which the optimization problems considered are convex. Furthermore, as mentioned in [30] and [31], where the problems are nonconvex, the convergence cannot be guaranteed.

C-2-2 Alternating Direction Method of Multipliers

As noted in [27], the Alternating Direction Method of Multipliers (ADMM) is one of the Augmented Lagrangian methods which were developed to improve the dual ascent method, in particular to yield convergence without assumptions such as strict convexity or finiteness of the cost function. Additionally, the authors of [28] also have observed that ADMM is able to improve convergence speed of the dual ascent in practice. In this section, a brief summary of this method is given. We refer the readers to [27] for a more extensive explanation of this method.

The following convex optimization problem is considered to explain the method:

$$\begin{aligned} \underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{minimize}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & A\boldsymbol{x} = \boldsymbol{c}, \end{aligned} \tag{C-30}$$

Algorithm 5 The Dual Ascent Method

Initialize $r = 1$ and $\boldsymbol{\lambda}^{(1)}$
while $\sum_{i=1}^{N} A_i \boldsymbol{x}_i^\star - c \geq \epsilon$ **do**
for for $i = \{1, 2, \cdots, N\}$ **do**
Solve the local optimization problems as indicated in using $\boldsymbol{\lambda}^{(q)}$ to obtain the current solution $\boldsymbol{x}_i^{(q)} \quad \forall i \in \{1, 2, \cdots, N\}$, that is (for each agent):

$$\boldsymbol{x}_i^{(q)} = \arg \min_{\boldsymbol{x}_i \in \mathcal{X}_i} \left(J_i(\boldsymbol{x}_i) + \boldsymbol{\lambda}^{(q)T} A_i \boldsymbol{x}_i \right).$$

end for
Update the Lagrange multiplier using the gradient ascent method:

$$\boldsymbol{\lambda}^{(q+1)} = \max \left(0, \boldsymbol{\lambda}^{(q)} + \gamma \left(\sum_{i=1}^{N} A_i \boldsymbol{x}_i^{(q)} - c \right) \right).$$

$r = r + 1$.
end while

where $A \in \mathbb{R}^{p \times n}$, $\boldsymbol{c} \in \mathbb{R}^p$, and $f : \mathbb{R}^n \to \mathbb{R}$ is convex. The augmented Lagrangian for this problem is

$$L_\rho(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^T (A\boldsymbol{x} - \boldsymbol{c}) + \frac{\rho}{2} ||A\boldsymbol{x} - \boldsymbol{c}||_2^2, \tag{C-31}$$

where ρ is called the penalty parameter. Minimizing $\boldsymbol{x}$ over the augmented Lagrangian can also be viewed as the following problem:

$$\begin{aligned} \underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{minimize}} \quad & f(\boldsymbol{x}) + \frac{\rho}{2} ||A\boldsymbol{x} - \boldsymbol{c}||_2^2 \\ \text{subject to} \quad & A\boldsymbol{x} = \boldsymbol{c}, \end{aligned} \tag{C-32}$$

which is equivalent to the original problem (C-30).

Applying dual ascent to the modified problem (C-32) yields the iterative algorithm which is written as

$$\boldsymbol{x}^{(q+1)} := \underset{\boldsymbol{x}}{\text{argmin}} L_\rho(\boldsymbol{x}, \boldsymbol{\lambda}^{(q)}) \tag{C-33}$$

$$\boldsymbol{\lambda}^{(q+1)} := \boldsymbol{\lambda}^{(q)} + \rho(A\boldsymbol{x}^{(q+1)} - \boldsymbol{c}). \tag{C-34}$$

This algorithm is known as the method of multipliers. The existence of the penalty term in the augmented Lagrangian provides an advantage that the strong assumptions for the dual ascent method to converge as have been mentioned are not required. However, this advantage comes at the cost that when the cost function is separable, the augmented Lagrangian L_ρ is not separable [27]. This implies that the minimization of $\boldsymbol{x}$ in step (C-33) can not be carried out separately in paralel for each x_i.

ADMM is intended to improve the method multipliers so that it can be used to decompose a problem into smaller subproblems. Thus ADMM is suitable to be a distributed optimization algorithm while having the convergence properties of the method of multipliers. The algorithm

deals with the following problem

$$\begin{aligned} \underset{\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n}{\text{minimize}} \quad & f(\boldsymbol{x}) + g(\boldsymbol{y}) \\ \text{subject to} \quad & A\boldsymbol{x} + B\boldsymbol{y} = \boldsymbol{c}, \end{aligned} \tag{C-35}$$

where $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{y}\in \mathbb{R}^m$, $A \in \mathbb{R}^{p\times n}$, $A \in \mathbb{R}^{p\times m}$ $\boldsymbol{c} \in \mathbb{R}^p$. Moreover it is assumed that f and g are convex. We can say that Problem (C-35) is formed by splitting the decision variable of Problem (C-30) which is denoted by $\boldsymbol{x}$ there into two parts which are denoted by $\boldsymbol{x}$ and $\boldsymbol{y}$ here. This yields to splitting of the cost function and the matrix on the constraint as well.

The augmented Lagrangian for this problem is

$$L_\rho(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + g(\boldsymbol{y}) + \boldsymbol{\lambda}^T(A\boldsymbol{x} + B\boldsymbol{y} - \boldsymbol{c}) + \frac{\rho}{2}||A\boldsymbol{x} + B\boldsymbol{y} - \boldsymbol{c}||_2^2, \tag{C-36}$$

and ADMM consists of the following iterations:

$$\boldsymbol{x}^{(q+1)} := \underset{\boldsymbol{x}}{\text{argmin}} L_\rho(\boldsymbol{x}, \boldsymbol{y}^{(q)}, \boldsymbol{\lambda}^{(q)}), \tag{C-37}$$

$$\boldsymbol{y}^{(q+1)} := \underset{\boldsymbol{y}}{\text{argmin}} L_\rho(\boldsymbol{x}^{(q+1)}, \boldsymbol{y}, \boldsymbol{\lambda}^{(q)}), \tag{C-38}$$

$$\boldsymbol{\lambda}^{(q+1)} := \boldsymbol{\lambda}^{(q)} + \rho(A\boldsymbol{x}^{(q+1)} + B\boldsymbol{y}^{(q+1)} - \boldsymbol{c}), \tag{C-39}$$

where $\rho > 0$.

It is also discussed in [27] that very mild assumptions are required to guarantee the convergence of ADMM. They are: (1) The extended-real-valued functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $g : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ are closed, proper, and convex, and (2) The unaugmented Lagrangian has a saddle point. Under these assumptions, the ADMM algorithm satisfies residual convergence, that is the solution approaches feasibility ($(A\boldsymbol{x}^{(q+1)} + B\boldsymbol{y}^{(q+1)} - \boldsymbol{c}) \to 0$ as $r \to \infty$), objective convergence that is the cost function approaches the optimal value ($f(\boldsymbol{x}^{(q)}) + g(\boldsymbol{y}^{(q)}) \to p^\star$ as $q \to \infty$) and dual variable convergence ($\boldsymbol{\lambda}^{(q)} \to \boldsymbol{\lambda}^\star$ as $q \to \infty$).

When dealing with a generic constrained optimization problem, ADMM can be formulated by first introducing an indicator function of the set that constrains the problem and an auxiliary variable which is associated to the set. The update of the auxiliary variable is then done by using a projection operator. As extracted from [27], This method is described as the following.

Consider a generic constrained convex optimization problem as follows

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & x \in \mathcal{P}, \end{aligned} \tag{C-40}$$

where f and $\mathcal{P}$ are convex. This problem can be rewritten in ADMM form (C-35) as

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & f(\boldsymbol{x}) + g(\boldsymbol{y}) \\ \text{subject to} \quad & \boldsymbol{x} - \boldsymbol{y} = 0, \end{aligned} \tag{C-41}$$

where g is the indicator function of $\mathcal{P}$, that is $g(\boldsymbol{y}) = 0$ if $y \in \mathcal{P}$ and $g(\boldsymbol{y}) = +\infty$ otherwise.

The augmented Lagrangian is

$$L_\rho(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + g(\boldsymbol{y}) + \boldsymbol{\lambda}^T(\boldsymbol{x} - \boldsymbol{y}) + \frac{\rho}{2}||\boldsymbol{x} - \boldsymbol{y}||, \tag{C-42}$$

so the ADMM for this problem is:

$$\begin{aligned}\boldsymbol{x}^{(q+1)} &:= \underset{\boldsymbol{x}}{\operatorname{argmin}} L_\rho(\boldsymbol{x}, \boldsymbol{y}^{(q)}, \boldsymbol{\lambda}^{(q)}) \\ \boldsymbol{y}^{(q+1)} &:= \Pi_{\mathcal{P}}\left(\boldsymbol{x}^{(q+1)}\right) \\ \boldsymbol{\lambda}^{(q+1)} &:= \boldsymbol{\lambda}^{(q)} + \rho\left(\boldsymbol{x}^{(q+1)} - \boldsymbol{y}^{(q+1)}\right).\end{aligned}$$

Note that $\Pi_{\mathcal{P}}(\cdot)$ is the projection operator onto the set $\mathcal{P}$.

It is a general ADMM formulation which is used as the basis of the decomposition method for an MPC problem. It can be seen in the work of [29] in which a distributed MPC algorithm based on ADMM is proposed. Moreover, as discussed in [27], it could also deal with consensus problems that inspired the work of [28] in which the authors provide a distributed model predictive consensus algorithm.

The ADMM approach to solve the optimal exchange problem which is related to our study is also worth to be mentioned. As discussed in [27], the exchange problem is given as follows

$$\begin{aligned}\underset{\{\boldsymbol{x}_i \in \mathbb{R}^n\}_{i=1}^N}{\text{minimize}} \quad & \sum_{i=1}^{N} f_i(\boldsymbol{x}_i) \\ \text{subject to} \quad & \sum_{i=1}^{N} \boldsymbol{x}_i = 0,\end{aligned} \tag{C-43}$$

where f_i denotes the cost function for subsystem i. The ADMM for this problem consists of the following iterations:

$$\boldsymbol{x}_i^{(q+1)} = \underset{\boldsymbol{x}_i}{\operatorname{argmin}} \quad f_i(\boldsymbol{x}_i) + \boldsymbol{\lambda}^{(q)T}(\boldsymbol{x}_i) + \frac{\rho}{2}\left|\left|\boldsymbol{x}_i + \left(\bar{\boldsymbol{x}}^{(q)} - \boldsymbol{x}_i^{(q)}\right)\right|\right|_2^2, \tag{C-44}$$

$$\bar{\boldsymbol{x}}^{(q+1)} = \sum_{i=1}^{N} \boldsymbol{x}_i^{(q+1)}, \tag{C-45}$$

$$\boldsymbol{\lambda}^{(q+1)} = \boldsymbol{\lambda}^{(q)} + \rho\bar{\boldsymbol{x}}^{(q+1)}. \tag{C-46}$$

Bibliography

[1] V. Rostampour and T. Keviczky, "Robust randomized model predictive control for energy balance in smart thermal grids," *European Control Conference*, 2016, to be published.

[2] A. Eisentraut and B. Adam, "Heating without global warming," *Featured Insight*, pp. 1–92, 2014.

[3] G. C. Calafiore and M. C. Campi, "The scenario approach to robust control design," *IEEE Transaction on Automatic Control*, vol. 51, no. 5, pp. 742–753, 2006.

[4] H. Lund, S. Werner, R. Wiltshire, S. Svendsen, J. E. Thorsen, F. Hvelplund, and B. V. Mathiesen, "4th Generation District Heating (4GDH). Integrating smart thermal grids into future sustainable energy systems.," *Energy*, vol. 68, pp. 1–11, 2014.

[5] P. Pflaum, M. Alamir, and M. Y. Lamoudi, "Comparison of a primal and a dual decomposition for distributed MPC in smart districts," *IEEE International Conference on Smart Grid Communications*, pp. 55–60, 2014.

[6] G. Chicco and P. Mancarella, "Distributed multi-generation: A comprehensive view," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 3, pp. 535–551, 2009.

[7] D. Wu and R. Wang, "Combined cooling, heating and power: A review," *Progress in Energy and Combustion Science*, vol. 32, no. 5-6, pp. 459–495, 2006.

[8] P. D. Christofides, R. Scattolini, D. Muñoz de la Peña, and J. Liu, "Distributed model predictive control: A tutorial review and future research directions," *Computers & Chemical Engineering*, vol. 51, pp. 21–41, 2013.

[9] G. C. Calafiore, "Random convex programs," *SIAM Journal on Optimization*, vol. 20, pp. 3427–3464, 2010.

[10] G. Schildbach, G. C. Calafiore, L. Fagiano, and M. Morari, "Randomized model predictive control for stochastic linear systems," *American Control Conference*, pp. 417–422, 2012.

[11] M. C. Campi, S. Garatti, and M. Prandini, "The scenario approach for systems and control design," *Annual Reviews in Control*, vol. 33, no. 2, pp. 149–157, 2009.

[12] J. Matusko and F. Borrelli, "Scenario-based approach to stochastic linear predictive control," *51st IEEE Conference on Decision and Control*, pp. 5194–5199, 2012.

[13] M. Prandini, S. Garatti, and J. Lygeros, "A randomized approach to stochastic model predictive control," *51st IEEE Conference on Decision and Control (CDC)*, pp. 7315–7320, 2012.

[14] Y. Long, S. Liu, L. Xie, and K. Johansson, "A scenario-based distributed stochastic MPC for building temperature regulation," *IEEE International Conference on Automation Science and Engineering (CASE)*, pp. 1091–1096, 2014.

[15] V. Rostampour, P. M. Esfahani, and T. Keviczky, "Stochastic nonlinear model predictive control of an uncertain batch polymerization reactor," *IFAC Conference on Nonlinear Model Predictive Control*, pp. 540–545, 2015.

[16] V. Rostampour, K. Margellos, M. Vrakopoulou, M. Prandini, and J. Lygeros, "Reserve requirements in AC power systems with uncertain generation," *4th IEEE PES Innovative Smart Grid Technologies Europe*, pp. 1–5, 2013.

[17] G. C. Calafiore, D. Lyons, and L. Fagiano, "On mixed-integer random convex programs," *51st IEEE Conference on Decision and Control (CDC)*, pp. 3508–3513, 2012.

[18] K. Margellos, V. Rostampour, M. Vrakopoulou, M. Prandini, G. Andersson, and J. Lygeros, "Stochastic unit commitment and reserve scheduling: A tractable formulation with probabilistic certificates," *European Control Conference*, pp. 2513–2518, 2013.

[19] K. Margellos, P. Goulart, and J. Lygeros, "On the road between robust optimization and the scenario approach for chance constrained optimization problems," *IEEE Transactions on Automatic Control*, vol. 59, no. 8, pp. 2258–2263, 2014.

[20] R. Negenborn and J. Maestre, "Distributed model predictive control: An overview and roadmap of future research opportunities," *IEEE Control Systems*, vol. 34, pp. 530–535, 2014.

[21] J. M. Maestre and R. R. Negenborn, eds., *Distributed model predictive control made easy.* Sevilla and Delft: Springer, 2014.

[22] P. Giselsson, "Distributed Model Predictive Control with suboptimality and stability guarantees," *49th IEEE Conference on Decision and Control (CDC)*, pp. 7272–7277, 2010.

[23] M. D. Doan, T. Keviczky, and B. De Schutter, "A distributed optimization-based approach for hierarchical MPC of large-scale systems with coupled dynamics and constraints," *IEEE Conference on Decision and Control and European Control Conference*, vol. 1, no. 4, pp. 5236–5241, 2011.

[24] P. Giselsson, M. D. Doan, T. Keviczky, B. D. Schutter, and A. Rantzer, "Accelerated gradient methods and dual decomposition in distributed model predictive control," *Automatica*, vol. 49, no. 3, pp. 829–833, 2013.

[25] R. Cheng, F. Forbes, and W. S. Yip, "Price-driven coordination for solving plant-wide MPC problems," *Journal of Process Control*, vol. 17, pp. 429–438, 2007.

[26] R. Martí, D. Sarabia, D. Navia, and C. De Prada, "Coordination of distributed model predictive controllers using price-driven coordination and sensitivity analysis," *10th IFAC International Symposium on Dynamics and Control of Process Systems*, pp. 215–220, 2013.

[27] S. Boyd, N. Parikh, B. P. E Chu, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2010.

[28] T. H. Summers and J. Lygeros, "Distributed model predictive consensus via the alternating direction method of multipliers," *50th Annual Allerton Conference on Communication, Control and Computing (Allerton)*, no. 1, pp. 79–84, 2012.

[29] F. Farokhi, I. Shames, and K. H. Johansson, "Distributed MPC Via dual decomposition and alternating direction method of multipliers," *Intelligent Systems, Control and Automation: Science and Engineering*, vol. 69, pp. 115–131, 2014.

[30] G. K. H. Larsen, N. D. Van Foreest, and J. M. a. Scherpen, "Distributed control of the power supply-demand balance," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 828–836, 2013.

[31] G. K. H. Larsen, N. D. V. Foreest, and J. M. A. Scherpen, "Distributed MPC applied to a network of households with micro-CHP and heat storage," vol. 5, no. 4, pp. 2106–2114, 2014.

[32] C. R. Touretzky and M. Baldea, "Integrating scheduling and control for economic MPC of buildings with energy storage," *Journal of Process Control*, vol. 24, no. 8, pp. 1292–1300, 2014.

[33] A. Lefort, R. Bourdais, G. Ansanay-alex, and H. Guéguen, "Hierarchical control method applied to energy management of a residential house," *Energy & Buildings*, vol. 64, pp. 53–61, 2013.

[34] M. Ellis, M. Heidarinejad, and P. D. Christofides, "Economic model predictive control of nonlinear singularly perturbed systems," *Journal of Process Control*, vol. 23, no. 5, pp. 743–754, 2013.

[35] E. H. A. van Vliet, *Flexibility in heat demand at the TU Delft campus smart thermal grid with phase change materials.* Master's thesis, Delft University of Technology, 2013.

[36] C. M. Grinstead and J. L. Snell, "Markov Chain," in *Introduction to Probability*, ch. 11, pp. 405–470, Rhode Island: American Mathematical Society, 2nd ed., 1997.

[37] V. Rostampour, *Tractable Reserve Scheduling Formulations for Power Systems with Uncertain Generation.* Master's thesis, Politecnico di Milano, 2012.

[38] J. Lofberg, "YALMIP : A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, pp. 284–289, 2004.

[39] Gurobi Optimization Inc, "Gurobi Optimizer Reference Manual," 2015.

[40] B. Nordell, A. Snijders, and L. Stiles, "The use of aquifers as thermal energy storage (TES) systems," *Advances in Thermal Energy Storage Systems*, pp. 87–115, 2015.

[41] M. Bloemendal, T. Olsthoorn, and F. Boons, "How to achieve optimal and sustainable use of the subsurface for Aquifer Thermal Energy Storage," *Energy Policy*, pp. 1–11, 2013.

[42] V. Rostampour, M. Jaxa-Rozen, M. Bloemendal, and T. Keviczky, "Building climate energy management in smart thermal grids via aquifer thermal energy storage systems," *Journal of Energy Procedia, Elsevier*, 2016, submitted.

[43] V. Rostampour, M. Bloemendal, M. Jaxa-rozen, and T. Keviczky, "A control-oriented model for combined building climate comfort and aquifer thermal energy storage system," *European Geothermal Congress*, 2016, accepted.

[44] R. M. Freund, "Applied Lagrange duality for constrained optimization," *Massachussetts Institute of Technology*, 2004.

[45] S. Boyd and L. Vandenberghe, *Convex Optimization*, vol. 25. Cambridge university press, 2010.

[46] J. Zeng, D.-Y. Xue, and D.-c. Yuan, "Research and development trend of distributed MPC," *2008 Fourth International Conference on Natural Computation*, vol. 4, pp. 417–421, 2008.

Glossary

List of Acronyms

ADMM	Alternating Direction Method of Multipliers
ATES	Aquifer Thermal Energy Storage
BTES	Borehole Thermal Energy Storage
CHP	Combined Heat Power
LEA	Low Energy Architecture
MPC	Model Predictive Control

List of Symbols

α_g	Step size of the gradient method
$\alpha_{\mathbf{c}}$	Volume-to-energy coefficient of the cold well
$\alpha_{\mathbf{h}}$	Volume-to-energy coefficient of the warm well
β	Level of confidence
$\boldsymbol{\gamma}$	Bound vector of the uncertainty
$\boldsymbol{\lambda}$	Lagrange multiplier
Δ	Uncertain set
ΔT	Temperature difference [K]
ϵ	Stopping value of the gradient-based method
η	Efficiency
Γ	Matrix associated with the decision variables in the seasonal storage dynamics
Λ	Start-up cost
Φ	Matrix associated with the state in the seasonal storage dynamics
ρ	Penalty parameter of ADMM

ρ_a	Air density [kg/m^3]
ρ_{w}	Water density [kg/m^3]
ε	Level of violation
(q)	Iteration index
$\boldsymbol{J}$	Cost function
$\boldsymbol{u}$	Decision/manipulated variable
$\boldsymbol{v}$	Binary variable
$\boldsymbol{y}$	Auxiliary variable in ADMM formulations
$\boldsymbol{\psi}$	Indicator function
$\mathbb{P}$	Probability operator
$\mathbb{R}$	Set of real number
$\mathbb{R}_+$	Set of non-negative real number
$\mathcal{B}$	Bounded set of the uncertainty
$\mathcal{C}$	Set of coupling constraints
$\mathcal{L}$	Set of local constraints
$\mathcal{N}$	Set of the agents
$\mathcal{N}_{-i}$	Set of the neighbors
$\mathbf{E}$	Matrix associated with the real decision variables in the local constraints
$\mathbf{F}$	Matrix associated with the binary variables in the local constraints
$\mathbf{G}$	Matrix associated with the real decision variables in the coupling constraints
$\mathbf{P}$	Constant vector in the local constraints
A	Matrix related to the state in the state-space description
a	Area [m^2]
B	Matrix related to the input in the state-space description
c	Thermal energy (negative) [KWh]
c_{p}	Specific heat capacity
c_f	Convection factor
d	Direction of the gradient method
$d^\star$	Optimal value of the dual problem
e	Auxiliary decision variable in the local constraints of the supply levels of warm and cold wells
g_d	Lagrange dual function
H	Matrix related to the uncertain variable
h	Thermal energy (positive) [KWh]
K	Matrix associated with the decision variables in the coupling constraints of the ADMM with coordinator
k	Sampling time
L	The Lagrangian
m	Sampling time of the upper layer in the hierarchical MPC
N	Number of agents

N_d	Number of decision variables
N_h	Prediction horizon
N_s	Number of scenarios
p	Electrical energy [KWh]
$p^\star$	Optimal value of the primal problem
Q	Weight matrix of the state
q_{sol}	Solar radiation [KWh]
R	Weight matrix of the decision variables
r	Weight coefficient
s	Level of the supply of a seasonal storage
s	The index of the scenario
T	Temperature [K]
t	Prediction time
V	Volume [m^3]
w	Uncertain variable
x	State
z	Auxiliary variable of minimum on/off constraints
h	Thermal energy [KWh]

Index