

2013

# Topics in Demand Response for Energy Management in Smart Grid

Chen Chen  
*Lehigh University*

Follow this and additional works at: <http://preserve.lehigh.edu/etd>



Part of the [Electrical and Computer Engineering Commons](#)

---

## Recommended Citation

Chen, Chen, "Topics in Demand Response for Energy Management in Smart Grid" (2013). *Theses and Dissertations*. Paper 1454.

This Dissertation is brought to you for free and open access by Lehigh Preserve. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Lehigh Preserve. For more information, please contact [preserve@lehigh.edu](mailto:preserve@lehigh.edu).

TOPICS IN DEMAND RESPONSE FOR ENERGY  
MANAGEMENT IN SMART GRID

BY

CHEN CHEN

PRESENTED TO THE GRADUATE AND RESEARCH COMMITTEE  
OF LEHIGH UNIVERSITY  
IN CANDIDACY FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

IN  
ELECTRICAL ENGINEERING  
LEHIGH UNIVERSITY

MAY 2013

© Copyright by Chen Chen, 2013.

All Rights Reserved

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

---

Date

---

Dissertation Advisor

---

Accepted Date

Committee Members:

---

Dr. Shaline Kishore

---

Dr. Rick S. Blum

---

Dr. Parv Venkatasubramaniam

---

Dr. Lawrence V. Snyder

# Acknowledgements

Writing acknowledgments of this dissertation is like a process of summarizing my Ph.D. life at Lehigh University. I feel very lucky to meet many people here who provided me all kinds of helps during these four years at Lehigh.

Without doubt, my deepest gratitude goes to my Ph.D. advisor, Prof. Shalinee Kishore, for her guidance, support, and persistent encouragement during my Ph.D. studies. The open attitude of Prof. Kishore towards research has granted me freedom to explore interesting topics to work on. At the same time, the insightful advices and comments from her on my research are invaluable fortunes, not only to this dissertation, but also to my attitude towards research in the future. Moreover, I also deeply appreciate her foresighted advices on my career path.

I would like to thank Prof. Rick Blum, Prof. Parv Venkitasubramaniam and Prof. Lawrence Snyder for kindly serving on my Ph.D. committees and providing insightful advices, suggestions on my research work and this dissertation. Discussions with them during my Ph.D. studies and in my general examination really benefited me enormously. In addition, I owe my gratitude to Guenter Conzelmann, Dr. Jianhui Wang, Dr. Yeonsook Heo, and Dr. Cong Liu and other wonderful colleagues I worked with during my internship at Decision and Information Sciences Division, Argonne National Laboratory, and I especially thank Dr. Jianhui Wang, my supervisor at Argonne, for providing me an open and exciting environment to do the research; part of the work in this dissertation was initiated when I was there. I also would like to thank Prof. Anna Scaglione, Dr. Zhifang Wang and Mahnoosh Alizadeh from University of California Davis for collaboration on part of the work in this dissertation.

I am grateful to my colleagues and friends at Lehigh University who have offered helps in various ways. I thank my fellow graduate students in the group: Gang Xiong, K.G. Nagananda, Sara Truesdale for the numerous discussions we have had. I would also like to thank my friends at Lehigh: Feng Shi & Lihua Jiao, Yang Liu, Jiadong Wang & Fang Ma , Hongmei Xie, Jun Lin, Ziad Rawas, Yang Yang, Chenrong Xiong & Yan Yang, Zhengjun Chen, Xiong Xiong & Xiaocun Que, Jiangfan Zhang & Xuanxuan Lv, and a list that is probably too long to go through. With you all, I really enjoy my life at Lehigh.

Finally, I would like to give my special appreciation to my father, Guoliang Chen, and my mother, Guaixia Wu, for their selfless love and care to me all over my life. Especially when I was in difficult times, their endless supports and constant encouragements make me be fearless towards every challenge in my life. Without them, this dissertation can not be completed.

*Dedicated to my parents!*

# Contents

- Acknowledgements . . . . . iv
- List of Tables . . . . . xi
- List of Figures . . . . . xii
- Abstract . . . . . 1
  
- 1 Introduction . . . . . 4**
  - 1.1 Background . . . . . 4
    - 1.1.1 Demand response . . . . . 4
    - 1.1.2 Three levels of DR in the smart grid . . . . . 6
  - 1.2 Outline of the Dissertation . . . . . 11
  
- 2 Residential Appliance Scheduling in Buildings: An Optimization Approach . . . . . 15**
  - 2.1 Introduction . . . . . 15
  - 2.2 System Model . . . . . 17
  - 2.3 Operation Flexible Appliance Scheduling . . . . . 20
    - 2.3.1 Delay flexible appliance scheduling . . . . . 20
    - 2.3.2 PEV battery charging/discharging scheduling . . . . . 25
  - 2.4 Thermal Appliance Scheduling . . . . . 29
    - 2.4.1 Thermal dynamics of building room . . . . . 29
    - 2.4.2 PMV comfort level . . . . . 32

2.4.3	Optimization formulation . . . . .	33
2.5	Numerical Results . . . . .	36
2.6	Chapter Summary . . . . .	44
<b>3</b>	<b>Residential Appliance Scheduling: A Communication-based Control Approach</b>	<b>45</b>
3.1	Introduction . . . . .	45
3.2	Access and Scheduling Scheme for Appliances . . . . .	47
3.2.1	Protocol description . . . . .	47
3.2.2	Admission control methods (ACM) . . . . .	50
3.2.3	Critical appliance consideration . . . . .	52
3.3	Markov Chain Analysis of the Protocol . . . . .	52
3.3.1	Power usage statistic . . . . .	53
3.3.2	Two-dimensional discrete-time Markov chain model . . . . .	54
3.3.3	Steady-state distribution . . . . .	57
3.3.4	Delay analysis . . . . .	58
3.4	Optimization of $\mathcal{P}_{\max,t}$ for EMC . . . . .	60
3.4.1	Wind power statistics . . . . .	61
3.4.2	Optimization formulation . . . . .	62
3.5	Numerical Results . . . . .	64
3.6	Chapter Summary . . . . .	69
<b>4</b>	<b>Reducing Peak Demand by Real-time Pricing</b>	<b>70</b>
4.1	Introduction . . . . .	70
4.2	Stackelberg Game Model Analysis . . . . .	72
4.2.1	EMC/follower level decisions . . . . .	73
4.2.2	Service provider/leader level decisions . . . . .	74
4.2.3	Equilibrium of the Stackelberg game . . . . .	76

4.3	RTP-based Power Scheduling Scheme . . . . .	77
4.4	Numerical Results . . . . .	79
4.5	Chapter Summary . . . . .	81
<b>5</b>	<b>Matching Demand and Supply: A Distributed Direct Load Control Approach</b>	<b>82</b>
5.1	Introduction . . . . .	82
5.2	Problem Formulation . . . . .	86
5.3	Distributed Demand Target Allocation . . . . .	87
5.3.1	Allocating demand target ( $Z_\tau$ ) to EMCs . . . . .	87
5.3.2	Distributed consensus algorithm to compute $\theta_i^\tau$ . . . . .	89
5.4	Lower-Layer Communication and Admission Control Scheme . . . . .	93
5.4.1	Load information update . . . . .	95
5.4.2	Admission control mechanism . . . . .	97
5.4.3	Non-intrusive operation for appliances . . . . .	99
5.5	Numerical Results . . . . .	104
5.6	Chapter Summary . . . . .	106
<b>6</b>	<b>Market Effects of Deferrable Loads as Demand Response</b>	<b>107</b>
6.1	Introduction . . . . .	107
6.2	Modeling the Deferrable Load . . . . .	109
6.3	Cournot Game of Wholesale Market with DR Aggregator Participation	112
6.3.1	Optimization for generators . . . . .	113
6.3.2	Optimization for DR aggregator . . . . .	114
6.3.3	Optimization for the ISO . . . . .	116
6.4	Equilibrium Analysis . . . . .	117
6.4.1	Optimality conditions for DR aggregator . . . . .	118
6.4.2	Optimality conditions for the ISO and generators . . . . .	122

6.4.3	Existence and uniqueness of the equilibrium . . . . .	125
6.5	Numerical Results . . . . .	126
6.6	Chapter Summary . . . . .	129
	<b>Bibliography</b>	<b>130</b>
	<b>Vita</b>	<b>142</b>

# List of Tables

2.1	Parameters for PMV computation . . . . .	32
2.2	Pricing schemes comparison . . . . .	36
2.3	Average temperature comparison for two places . . . . .	37
2.4	Appliances (events) parameters . . . . .	37
2.5	Cost comparison for different pricing schemes . . . . .	40
4.1	Appliances' parameters for each home . . . . .	79
5.1	Fields description in appliance TX packet . . . . .	96
5.2	Fields description in EMC TX packet . . . . .	97

# List of Figures

1.1	Three levels of DR implementation in electricity market . . . . .	6
1.2	Two types of approaches for DR in retail market . . . . .	8
2.1	Residential building energy management system . . . . .	18
2.2	MPC-based appliance scheduling . . . . .	19
2.3	Building room prototype and model . . . . .	30
2.4	PMV data computed using iterative algorithm . . . . .	33
2.5	One-day scheduling results for pricing scheme 2 . . . . .	39
2.6	One-day battery scheduling result for pricing scheme 3 . . . . .	40
2.7	Energy cost comparison for winter and summer . . . . .	41
2.8	Impact of customers' preference . . . . .	42
2.9	Time horizon length trade-off under price uncertainty . . . . .	43
3.1	Frame structure of proposed protocol . . . . .	48
3.2	Admission success probability given $ \mathcal{D}_t  = q$ and $ \mathcal{A}_t  = n$ . . . . .	56
3.3	State transition diagram for an arbitrary appliance . . . . .	58
3.4	Transition probability matrix of wind power . . . . .	62
3.5	Average delay versus $\mathcal{P}_{\max}$ with various $(\lambda, \mu)$ (random ACM, $N = 15$ )	64
3.6	Average delay versus $N$ with various ACMs. ( $\mathcal{P}_{\max} = 3000\text{W}$ , $\lambda =$ 1/900, and $\mu = 1/3000$ ) . . . . .	65
3.7	Critical appliances effects. ( $N = 15$ , $\lambda = 1/900$ , and $\mu = 1/3000$ ) . .	66

3.8	Optimized power budget $\mathcal{P}_{\max,t}$ for a day . . . . .	67
3.9	Total power consumption under our scheme using $\mathcal{P}_{\max,t}$ with/without wind power. ( $N = 25$ , $\mu = 1/3600$ and random ACM) . . . . .	68
3.10	Comparison of electricity cost and average delay. . . . .	69
4.1	The interaction between consumers and electricity service provider .	73
4.2	One-day power usage comparison . . . . .	80
4.3	Money saving evaluation (100 days) . . . . .	81
5.1	Two-layer communication-based control structure . . . . .	83
5.2	Four sample networks with 20 nodes . . . . .	91
5.3	Convergence properties for four networks . . . . .	92
5.4	Two-layer communication and control protocol . . . . .	94
5.5	Aggregated demand profile for simulation . . . . .	105
5.6	Scheduling demand using the proposed approach . . . . .	106
6.1	Participation of DR aggregators in the wholesale market . . . . .	112
6.2	IEEE 24-bus test network with DR aggregators . . . . .	126
6.3	Cournot game price validation at node 20 (No DR aggregator) . . .	127
6.4	Price comparison with/without DR aggregator . . . . .	128
6.5	Price reduction and profit of DR aggregator . . . . .	129

# Abstract

Future electricity grids will enable greater and more sophisticated demand side participation, which refers to the inclusion of mechanisms that enable dynamic modification of electricity demand into the operations of the electricity market, known as *Demand Response* (DR). The underlying information-flow infrastructures provided by the emerging *smart grid* enhance the interactions between customers and the market, by which DR will improve electricity grids in several aspects, e.g., by reducing peak demand and reducing need for expensive peaker plants, or by enabling demand to follow supply such as those from volatile renewable resources, etc. Many types of appliances provide flexibilities in power usage which can be viewed as demand response resources, and how to exploit such flexibilities to achieve the benefits offered by DR is a central challenge. In this dissertation, we design algorithms and architectures to bridge the gap between scheduling appliances and the benefits that DR can bring to electricity grid by utilizing the smart grid's underlying information infrastructure.

First, we focus on demand response within the consumer premise, where an energy management controller (EMC) schedules appliance operation on behalf of customers to save energy cost. We propose an optimization-based control scheme for the EMC in the building that integrates both the operational flexible appliances such as clothes washer/dryer, dish washer and plug-in electric vehicles (PEVs), but also the thermostatically controlled appliances such as HVAC (heating, ventilation, and air conditioning) systems together with the thermal mass of the building. Model predictive control is employed to account for uncertainty in electricity prices and weather information. Under time-varying pricing, scheduling appliances smartly using our scheme can incur notable energy cost saving for customers. As

an alternative, we also propose a communication-based control approach which is a joint appliance access and scheduling scheme in which the control algorithms are embedded into the communication protocols used by appliances. The control scheme is based on a threshold maximum power consumption set by the EMC; and we discuss how this threshold can be chosen so that it integrates the availability of local distributed renewable energy resources.

Then we investigate demand response in the retail market level which involves interactions between customers and utilities. Pricing-based control and direct load control (DLC) are two types of approaches that are used or envisioned for this level. To address pricing based control methods, we propose real-time pricing (RTP) signals that can be designed to work with customer premise EMCs. The interaction between these EMCs and the pricing-setting utilities is modeled as a Stackelberg game. We demonstrate that our proposed RTP scheme reduces peak load and alleviates rebound peaks that are the typical shortcomings in existing pricing approaches. To address DLC methods, we propose a distributed DLC scheme based on a two-layer communication network infrastructure for large-scale, aggregate DR implementations. In the proposed scheme, average consensus algorithms are employed to distributively allocate control tasks amongst EMCs so that local appliance scheduling within each home will eventually achieve the aggregated control task, i.e., to alleviate mismatch between electricity supply and demand.

Finally, we study how demand response affects the wholesale electricity market. As is conventional when studying interactions between electricity generators, we employ the Cournot game model to analyze how DR aggregators may impact wholesale energy markets. To do so, we assume that DR aggregators employ a computationally efficient, centralized scheduling mechanism to manage deferrable load over a large aggregate set of consumers. The load reduction from deferrable load

can be seen as ‘generation’ in terms of balancing the market and is compensated as such under current regulatory mandates. Thus, the DR aggregator competes with other generators in a Cournot-Nash manner to make a profit in the wholesale market; and electricity prices are consequently reduced. We provide equilibrium analysis of the wholesale market that includes DR aggregators and demonstrate that under certain conditions the equilibrium exists and is unique.

# Chapter 1

## Introduction

### 1.1 Background

#### 1.1.1 Demand response

U.S. electricity grids face increasing pressure to match supply and demand. This pressure comes from two sources: On the one hand, electricity demand is increasing in quantity and changing in quality. For example, the Energy Information Administration (EIA) estimates that overall electricity demand will increase 30% from 3.9 trillion kWh in 2009 to 5.0 trillion kWh in 2035 [1]. A portion of this increase will come from wide adoption of electric vehicles (EVs) whose demand profile is significantly different from current loads. On the other hand, driven by increasing prices for fossil fuels and concerns about greenhouse gas (GHG) emissions, renewable energy resources are rapidly being introduced into the existing electricity supply portfolio. For example, the U.S. Department of Energy (DOE) estimates that wind power will meet 20% of the U.S. electricity demand by 2030,

which means that the U.S. wind power capacity may reach more than 300 gigawatts (GW) [2]. With emerging *smart grid* technologies, a key solution for the power grids to alleviate the stresses of increasing demand and intermittent renewable generation is to encourage *demand response* (DR) as *demand side participation* in the electricity market [3]. According to the DOE, DR is defined as

*“Changes in electric usage by end-use customers from their normal consumption patterns in response to changes in the price of electricity over time, or to incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized.”*

The capabilities of “changes in electric usage” in the definition above are the resources available for demand response. It is observed that certain types of appliances with operational flexibility may serve as candidate resources for DR. For example, some appliances offer flexibility in the operation time (e.g., dish washer or clothes washer/dryer) so that their start time can be delayed under the constraint of a hard deadline. The plug-in electric vehicle (PEV) is another important emerging load type that offers flexibility in the power consumed in charging PEV batteries. This type of load is a great DR asset because individual PEV battery charge can be modified (and coordinated across multiple PEV batteries) as long as its total energy requirement is fulfilled by a given deadline. Thus, DR schemes that exploit loads with flexible operation time and power draw can shift electricity usage in time and benefit both consumers and the grid [4]. From the grid’s perspective, shifting some load away from peak periods 1) lowers the cost of employing less efficient generation during peak demand; 2) alleviates transmission congestion; and 3) helps maintain grid stability during these critical peak hours. From the customer’s perspective, shifting load to off-peak periods results in reduced electricity bills and occurrences of blackouts and brownouts.

Due to the underlying two-way communication infrastructures enabled by the future smart grid renovation, the interactions between customers, utilities, and the market can be largely enhanced. A central challenge on implementing demand response in future smart grids is how to manage demand response resources, i.e., how to exploit the underlying information infrastructure and operational flexibilities of loads to achieve the promised benefits of DR.

### 1.1.2 Three levels of DR in the smart grid

Depending on the different levels of aggregation, DR implementation and participation in the layered structure of the electricity market can be divided into three categories as shown in Figure 1.1.

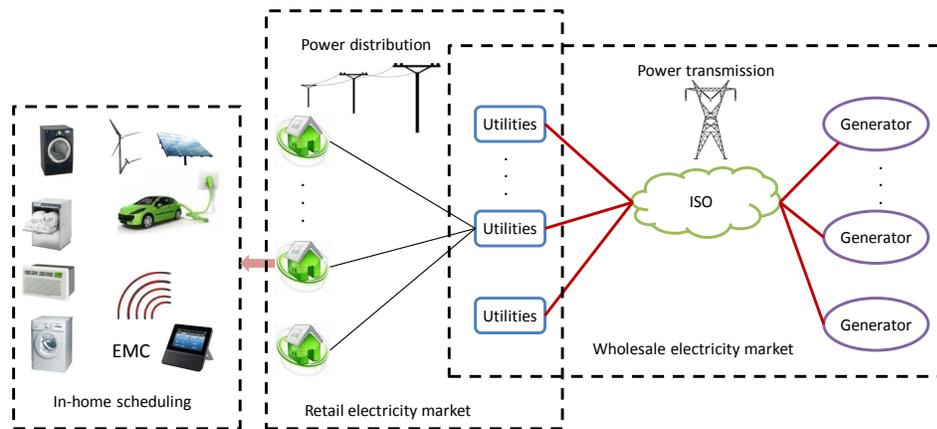


Figure 1.1: Three levels of DR implementation in electricity market

- Consumer premise level: The energy management controller (EMC) schedules operation of appliances within home to save money or fulfill the control task on behalf of customers;
- Retail market level: With appliance scheduling capabilities, the EMCs inter-

act with utilities in the retail electricity market;

- Wholesale market level: Large-scale aggregation of DR resources can participate in and affect the wholesale electricity market.

In this dissertation, we propose innovative algorithms and architectures for scheduling appliances as demand response in these three levels, and evaluate the benefits accordingly. The main contribution of our work is to bridge the gap between the scheduling of appliances and the benefits that demand response can bring to the grids utilizing underlying information infrastructure, and to guide the implementation of residential demand response scheme in future smart grids.

### **(1) DR in consumer premise level**

The demand response implemented in this level usually aims at minimizing the energy (electricity) cost by smartly scheduling appliances within a home. Besides the operational flexibilities provided by some appliances such as clothes washer/dryer, dish washer, thermal comfort flexibilities of customers provide additional freedom in scheduling thermostatically controlled appliances like HVAC (heating, ventilation and air conditioning) systems. In addition, several utilities, e.g., Ameren Illinois [5] and ComEd [6], already provide time-varying electricity tariffs to subscribed customers. Such customers can utilize their load flexibilities to save on electricity bills. However, current residential load control activities are mainly operated manually, which poses great challenges to customers in optimally scheduling the operations of their appliances. Customers may not have time to make such scheduling decisions and if prices vary fast and frequently, scheduling may be too complex. Hence, an automated energy management controller (EMC) is necessary to optimize the appliances' operation on behalf of customers.

In this dissertation, we design appliance scheduling schemes for the EMC to fully exploit the various degrees of flexibilities required by different appliances to save energy cost on behalf of customers. One approach is to formulate optimization problems of appliance scheduling in response to time-varying pricing, where operational and thermal flexibilities are integrated as a viable tool of changing the power usage profile to save money. An alternative approach is to combine the control and the underlying communication together, i.e., design a communication-based control scheme where the load control problem is embedded into the communication protocol.

## (2) DR in retail market level

The interactions between customers and utilities are usually involved when DR is implemented in the retail market level. By these interactions, utilities can make customers change their power usage profile, either directly or indirectly, to reduce peak demand and alleviate the mismatch between supply and demand. In general, there are two approaches for DR in the retail market, as shown in Figure 1.2:

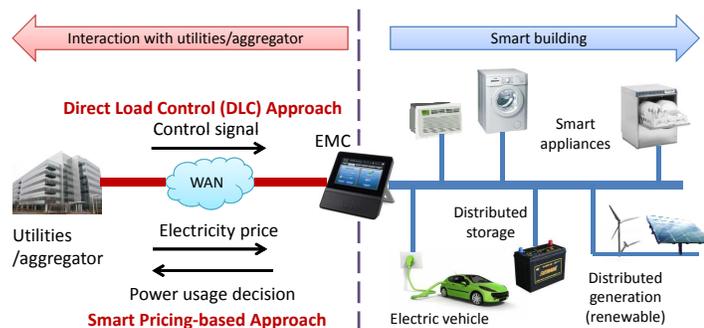


Figure 1.2: Two types of approaches for DR in retail market

*a. Direct load control (DLC) approaches:* Current DLC programs provided by utilities are usually contract-based, i.e., by signing up for the contract, customers

give utilities the option to remotely shut down appliances during high-demand periods or power supply emergency, and receive credit on electricity bills for this participation. Examples of these DLC programs include Contracted Direct Load Control by Wisconsin Public Service [7] and Distribution Load Relief Program by conEdison [8].

Several limitations exist for these types of DLC programs. Firstly, they are only for emergency cases so that they do not fully exploit the operational flexibilities of appliances which have potential for balancing supply and demand. Secondly, in scenarios with a large number of customers (and loads) participating, complexities of computation and communication are huge for the central controller. Another limit is due to customers' privacy concerns since their power usage is exposed as each individual appliance is remotely controlled by the central controller.

In this dissertation, we propose a distributed direct load control scheme which overcomes the drawbacks of the current DLC programs listed above. The scheduling of appliances is conducted within each home by the EMC, and all EMCs are coordinated by local communications so that the control task can be allocated in a distributed way.

*b. Smart pricing-based approaches:* Utilities can also choose to control appliances indirectly by sending price signals. Specifically, the variation of wholesale electricity prices is brought to the retail electricity prices and affects customers' electricity bills. With time-varying electricity prices, customers may be induced to make smart energy decisions, e.g., shifting some power usage during high-price periods to low-price periods which will save customers money and reduce overall burden on the grid during peak demand periods. Examples of implementation of these types of approaches include Ameren Illinois [5], and ComEd [6].

Although there are time-varying pricing schemes currently, none of them are fully dynamic. This means that there exists potential for ‘rebound peak demand’ [9], which are instances in which a significant number of customers shift their loads to low-price periods and cause a new peak. In some cases, this rebound peak can be far more problematic (i.e., far more of a mismatch to the available power supply) than the original high-price peak. A possible solution is to introduce real-time pricing (RTP) which reflects the more dynamical relationship between supply and demand in the market. In this dissertation, we propose an innovative RTP scheme which combines residential flexible appliance scheduling. The interaction between the utilities and customers is modeled as a Stackelberg game and the algorithms can be implemented via the underlying two-way communication network.

### **(3) DR in wholesale market level**

Depending on the Independent System Operator’s (ISO’s) market design, DR may provide energy, reserve ancillary service (A/S), or capacity in the wholesale electricity market [10]. Currently, several trials of DR programs have been implemented by some ISOs in the U.S., e.g. New York ISO (NYISO) [11], PJM [12] and ISO New England (ISO-NE) [13]. At the same time, stimulus of DR in the wholesale electricity market is also provided by regulatory institutions like Federal Energy Regulatory Commission (FERC). For example, FERC Order No.719 [14] and No. 745 [15] specify how ISOs can permit Demand Response Providers (DRPs) to bid DR on behalf of retail customers directly into the ISO’s organized markets, and get compensated for the service they provide at the locational marginal price (LMP). These efforts help integrate DR resources into the wholesale electricity market to cope with future supply-demand matching challenges.

When DR resources participate in the wholesale electricity market, aggregation is usually needed. In such cases, an important question for future DR systems is how will a DR aggregator manage operational flexibilities of loads across a large number of customers. The approaches we propose for DR in the retail market, i.e., for direct load control, could be one way to do this. In this dissertation, we build on an alternate computationally efficient architecture of managing deferrable loads as a market asset and employ Cournot game model to analyze the effects of DR aggregators' participation in the wholesale market.

## 1.2 Outline of the Dissertation

As previously mentioned, this dissertation mainly focuses on three levels of demand response in the grid: consumer-premise level DR where the EMC schedules appliances to save money on behalf of customer; retail market level DR where EMCs interact with the utilities to reduce peak demand or match the supply and demand; wholesale market level DR where the aggregators apply DR resources in the wholesale electricity market and influence the market, e.g., the electricity prices. A brief outline of this dissertation is presented as follows.

In Chapter 2, we apply model predictive control (MPC) method to appliance scheduling within a building under a time-varying pricing scenario. Specifically, the building thermal mass is integrated by modeling the building's thermal dynamics as linear difference equality constraints in its optimization. Thus, electricity cost can be saved when the scheduler utilizes the thermal mass (serving as heat storage) in conjunction with time-varying prices and weather information when scheduling loads. In addition, scheduling of non-thermal flexible appliances is modeled as a

mixed-integer linear program and both inter-appliance and intra-appliance dependencies are incorporated when scheduling in the face of time-varying prices. With the proposed scheme, customers are shown to have notably reduced electricity costs.

In Chapter 3, we focus on how a consumer premise communication network (e.g., a home-area network) can be designed to permit load scheduling amongst flexible and controllable loads. Our proposed scheme alleviates the EMC of the communication and computation complexities of directly scheduling each individual load in the consumer premise. The functionalities of the EMC are reduced in our model to calculating a threshold maximum power consumption for the home. We cast the calculation of this time-varying threshold power as an optimization problem which accounts for price variation and uncertainty due to local wind generation. Our proposed joint access and scheduling protocol describes how appliances access a common control channel so that this total maximum demand target is not exceeded for each time slot. Unlike existing local area network (LAN) media access protocols (e.g., Wi-Fi, G.hn, Zigbee), which could be used for one of several applications, our proposed scheme specifically addresses the load control problem. For in-home scenarios, relatively short control messages are occasionally exchanged; thus the approach can be easily implemented on higher data-rate LANs. To provide analytical foundations, we study the evolution of the protocol as a two-dimensional Markov chain and quantify the average delay experienced by individual appliances.

In Chapter 4, we move our focus on how the real-time pricing (RTP) can make a connection between scheduling appliances and the benefits of this demand response, e.g., peak load reduction. A Stackelberg game model is formulated to analyze the interaction between a customer's EMC and the serving electric utility,

where the utility plays the leader level game and the customer plays the follower level game. The RTPs are signals that can indirectly make customers change their power usage profile, and as a result, the peak load as well as the mismatch between supply and demand can be alleviated. Due to the dynamic adjusting of the pricing introduced in our approach, the ‘rebound peak demand’ can also be alleviated.

In Chapter 5, we aim at the direct load control (DLC) as an alternative approach for demand response in the retail market level. We propose a distributed direct load control approach based on a two-layer communication network infrastructure. The lower-layer network is within each building, where the energy management controller (EMC) uses wireless links to schedule operation of appliances upon request according to a local power consumption target. The upper-layer network links a number of EMCs in a region whose aggregated demand is served by a load aggregator. The load aggregator wants the actual (day-of) aggregated demand over this region to match a desired aggregated demand profile (i.e., day-ahead planned supply). Our approach utilizes the average consensus algorithm to distribute portions of the desired aggregated demand to each EMC in a decentralized fashion. The allocated portion corresponds to each building’s local power consumption target which its EMC then uses to schedule the in-building appliances. The result will be an aggregated demand over this region that more closely reaches the desired demand.

In Chapter 6, we investigate the impact of demand response at the wholesale market level. A Cournot game model is proposed to analyze the market effects with participation of DR aggregators, which manage aggregated deferrable load using a computationally efficient architecture. According to the FERC rules, the load reduction of the DR aggregator from deferrable loads can be viewed as ‘generation’ in terms of balancing the market. Thus the DR aggregator competes with other

traditional generators in a Cournot-Nash manner to make a profit in the market, and market prices are reduced as a consequence. With the Cournot game model, we provide equilibrium analysis of the wholesale market that includes the DR aggregators. We show that under certain conditions, the market equilibrium exists and is unique.

## Chapter 2

# Residential Appliance Scheduling in Buildings: An Optimization Approach

### 2.1 Introduction

In this chapter, we focus on demand response in residential buildings, and formulate optimizations for scheduling of several types of appliances. Our approach takes advantage of future two-way communication infrastructure underlying the emerging smart grid, and utilizes time-varying prices to save money for customers. The proposed optimization-based control algorithms can be integrated into a building energy management controller (BEMC) that schedules appliances on behalf of the customer automatically. The residential appliances for scheduling can be divided into non-thermal appliances with flexible operation (e.g. clothes washer/dryer, dish washer, plug-in electric vehicle (PEV)), and thermostatically controlled appliances

in which customers' comfort ranges provide a degree of operational flexibility.

For non-thermal flexible appliances scheduling, we formulate a mixed-integer linear programming (MIP) problem that captures the key features of appliance operations, and also the inter-appliance and intra-appliance dependences to exploit finer granularity when faced with time-varying prices. The MIP problems can be solved efficiently within the building scale, and can be easily embedded into the BEMC.

For thermostatically controlled appliances scheduling, our approach integrates the thermal mass of the building in a linear programming (LP) problem. Specifically, a simplified 14-node thermal dynamics model is proposed as linear difference equality constraints in the LP, such that the thermal mass of the building can represent heat storage and be effectively utilized to save money for customers. Model predictive control (MPC) is applied in the optimization to incorporate updated external information such as prices and weather information.

There are several studies about residential appliance scheduling. Certain types of delay flexible appliance scheduling optimization is formulated in [9] and [16]. The ON/OFF operations of individual appliances are modeled as independent Markov chains in [17]. However, other operation constraints like operational dependencies and minimum operation duration are omitted in these papers. For thermal appliances scheduling, Lu et al. presented a cost minimization problem for electric water heater (EWH) scheduling considering the thermal dynamics of water heating systems and customers' preferences [18]. The thermal dynamics of heating, ventilation, and air conditioning (HVAC) is considered in [19]. But the thermal system only incorporated the heat transfer between the room air and the outside environment, omitting the thermal mass storage of the building itself. An exponential

decay building thermal model is utilized in [20], and dynamic programming is used for thermal appliance scheduling. Instead of using an optimization formulation, simulation tests for certain pre-cooling strategies considering the thermal mass are conducted in [21]. An MPC method for thermal appliance scheduling considering the thermal dynamics of the building is employed in both [22] and [23]. The thermal model in [22] only considers a simple three-node case to represent the building while the thermal model used in [23] has non-linearity issues. Besides, both papers apply only the temperature as a thermal comfort indicator in their formulations.

The remainder of this chapter is organized as follows: Section 2.2 describes the system model. Section 2.3 formulates the optimization problems for non-thermal appliances scheduling. Section 2.4 incorporates the thermal mass of buildings in the thermal appliances scheduling optimization problem. Section 2.5 presents numerical simulation results and we conclude our discussion in Section 2.6.

## 2.2 System Model

In this paper, we consider a residential building, in which home appliances are scheduled by the BEMC, shown in Figure 2.1. In our model, appliances in the building can be divided into the following categories for scheduling:

- Delay flexible appliances, e.g., clothes washer/dryer, dish washer, etc.
- Delay and power consumption flexible appliances, such as a PEV battery.
- Thermal appliances, which include electric heater and A/C.

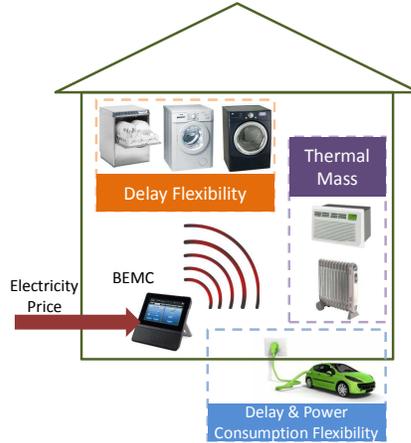


Figure 2.1: Residential building energy management system

The detailed models of these types of appliances for the scheduling optimization will be discussed in Sections 2.3 and 2.4. The BEMC schedules operation of appliances by exploiting these flexibilities. In our model, the BEMC (acting on behalf of the customer) is a price taker so that its scheduling decision does not affect electricity prices.

Model predictive control (MPC) refers to a class of control methods that compute a sequence of decision variable adjustments over a future time horizon iteratively based on an underlying optimization model and forecasts of uncertain variables. In other words, MPC is a rolling process that runs the embedded optimization model repeatedly with updated forecasts. MPC is now recognized as a very powerful approach with well established theoretical foundation and proven capability to handle a large number of control problems with uncertainty [22]. The thermal dynamics of the building which is viewed as a process model underlying the appliance scheduling optimization, and the prediction of time-varying prices as well as weather can all be incorporated in the MPC method for the BEMC to utilize the building thermal mass and time-varying pricing.

The MPC method is employed by the BEMC in scheduling appliances, shown in Figure 2.2. The time-line is divided into slots for scheduling, each time slot with a duration of  $\Delta$  minutes. The MPC-based approach works as follows: At any time slot  $t$ , the BEMC gets the current electricity price and weather data and forecast data from  $t + 1$  to  $t + N - 1$ . The BEMC then solves an optimization problem to minimize the energy cost of thermal appliances (A/C or heater) over this  $N$ -slot time horizon. The actual operation (power consumption) at time  $t$  will follow the optimization result for that time. Then at time  $t + 1$ , the BEMC obtains updated information for the next  $N$  time slots (from  $t + 1$  to  $t + N$ ) and solves this finite horizon optimization problem again, and schedules the power consumption for time  $t + 1$ . The time horizon moves forward by one time slot for the new optimization. For non-thermal flexible appliances scheduling, let  $\mathcal{R}_t$  denote the set of requested appliances at the beginning of  $t$  (these requests happen during  $t - 1$  slot), the BEMC optimizes the operation (ON/OFF status) or power consumptions for the next  $N$  slots starting from  $t$ , and the actual appliances' operation will follow the optimization results. As this finite time horizon optimization procedure schedules the appliances' operation considering not only the current information by also the forecast information, short-sightedness in scheduling can be avoided. Detailed optimization formulations are described in the next two sections, where  $\tau = 0, 1, \dots, N - 1$  is used for the relative time index within the time horizon  $N$ .

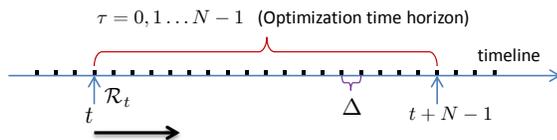


Figure 2.2: MPC-based appliance scheduling

Note that the future information, i.e., electricity prices and weather data, within the time horizon  $N$  need forecasting techniques, which is not focus of our approach.

Instead, we assume that the BEMC obtains perfect forecasts, which use the real prices and weather data in the simulations as shown in Section 2.5. Extensive studies about electricity price forecasting can be found in [24] [25], etc.

## 2.3 Operation Flexible Appliance Scheduling

In this section, the optimization problem of scheduling non-thermal appliances is formulated. We consider two categories of appliances: delay flexible appliances (such as dishwashers, clothes washers and clothes dryers) and both delay and power consumption flexible appliances (such as PEV batteries).

In the first category, rather than scheduling each device, the BEMC schedules the *task* of a user's power consumption *event*, which considers both *intra-appliance* and *inter-appliance* dependencies. In the second category, we consider both charging and discharging modes of a PEV battery so that it can further help to save energy cost.

### 2.3.1 Delay flexible appliance scheduling

In a household, the operation of delay flexible appliances usually has the following features: 1) The power consumption operation depends on the customers' request, i.e., customers pushing the start button of the appliance. 2) There is interdependence in appliances' operation, which can be either *intra-appliance* or *inter-appliance*. Intra-appliance dependence refers to cycles of an appliance operation. For example, the clothes washer has washing, rinsing, spinning, and maintenance wash cycles, where each cycle may run several times depending on the user's

configuration. Inter-appliance dependence refers to the interaction of different appliances. For example, a clothes dryer must operate after completion of a clothes washer. 3) Usually, some amount of time delay between tasks is allowed, for both intra-appliance and inter-appliance tasks.

We designed our flexible appliance scheduling algorithm according to the features listed above. When there is a power request from the customer, the scheduling algorithm will be triggered at the beginning of next nearest slot. Rather than scheduling the appliance as a consecutive block of power consumption, the BEMC treats each task individually; and thus, considers the *intra-appliance* and *inter-appliance* dependencies uniformly. In this way, the delay flexibility is not only from the delay time between the customer's request and actual appliance powering-on, but also from the time interval between tasks. Optimally allocating the task schedule and intervals between tasks may utilize low-price periods and hedge against high-price periods.

Let  $i$  denote the index of an event, and  $k$  denote the index of a task for a certain event. We define a binary variable  $u_{i,\tau}^k$  to indicate the ON/OFF status of an appliance with regard to task  $k$  for event  $i$  at time slot  $\tau$  within the time horizon, where  $u_{i,\tau}^k = 1$  indicates the ON status. For ease of formulation, we also define ancillary binary variables  $z_{i,\tau}^k$  and  $y_{i,\tau}^k$  to represent the start-up and shut-down actions, respectively, of task  $k$  for event  $i$  at time slot  $\tau$  within the time horizon, where  $z_{i,\tau}^k = 1$  and  $y_{i,\tau}^k = 1$  mean the start-up and shut-down actions.

**(1) Objective function:** Since each event operates independently and given the assumption that the BEMC is a price taker (i.e., the power consumption of one event does not influence the price of electricity, and thus, it does not affect the

decision of scheduling another event), we model the optimization for each event individually.

The objective of scheduling is to minimize the energy cost, which is the money paid for the power consumption of that event over the time horizon of  $N$  slots.

$$\min_{u_{i,\tau}^k, y_{i,\tau}^k, z_{i,\tau}^k} \sum_{k=1}^{ST_i} \sum_{\tau=0}^{N-1} u_{i,\tau}^k \cdot P_i^k \cdot \kappa_\tau, \quad (2.1)$$

where  $\kappa_\tau$  is the electricity price at time slot  $\tau$ , and  $P_i^k$  is the power consumption level of appliance  $i$  for task  $k$ .  $ST_i$  denotes the number of tasks in event  $i$ .

**(2) Coupling constraints of individual task ON/OFF status, start-up and shut-down indicator:** (2.2)-(2.4) formulates the logic relation between the ON/OFF status, start-up and shut-down indicators.

$$u_{i,\tau}^k - u_{i,\tau-1}^k = z_{i,\tau}^k - y_{i,\tau}^k, \quad \forall i, k \in \{1..ST_i\}, \tau. \quad (2.2)$$

$$z_{i,\tau}^k \leq u_{i,\tau}^k, \quad \forall i, k \in \{1..ST_i\}, \tau. \quad (2.3)$$

$$y_{i,\tau}^k \leq 1 - u_{i,\tau}^k, \quad \forall i, k \in \{1..ST_i\}, \tau. \quad (2.4)$$

$$u_{i,\tau}^k, y_{i,\tau}^k, z_{i,\tau}^k \in \{0, 1\}, \quad \forall i, k \in \{1..ST_i\}, \tau. \quad (2.5)$$

**(3) Start-up and shut-down time constraints:** We assume each task of the event only operates once during the time horizon  $N$ , i.e., there is only one start-up and shut-down action of each task. If one cycle of operation happens several times during the event, these cycles are treated as different tasks sequentially. We write

these constraints as

$$\sum_{\tau=0}^{N-1} z_{i,\tau}^k = 1, \quad \sum_{\tau=0}^{N-1} y_{i,\tau}^k = 1, \quad \forall i, k \in \{1..ST_i\}. \quad (2.6)$$

**(4) Operation time constraints:** We assume that once a task has powered on, the corresponding appliance remains on for  $T_i^{k,\text{ON}}$  slots with a constant power consumption level of  $P_i^k$ . To map this requirement to the decision variable, we employ the minimum-ON constraints of the unit commitment problem [26] in the following.

$$\sum_{\tau'=\max(0,\tau-T_i^{k,\text{ON}}+1)}^{\tau} z_{i,\tau'}^k \leq u_{i,\tau}^k, \quad \forall \tau, i, k. \quad (2.7)$$

The inequality constraints in (2.7) guarantee that the operation time of task  $k$  is greater than  $T_i^{k,\text{ON}}$ . If electricity prices are positive, from the objective function (2.1), we see that operation time of  $T_i^{k,\text{ON}}$  always incur the lowest cost. So these constraints actually specify that the operation time of task  $k$  is equal to  $T_i^{k,\text{ON}}$ . For the case of negative electricity prices, the task operation time may be extended to earn extra money. If tasks allow extra operation time, our scheme still applies; for the tasks with strict operation time, a hard deadline can be set in the actual scheduling.

**(5) Operation dependence constraints:** Without loss of generality, we assume the operation order of tasks for an event to be ascending in the index, i.e., task  $k$  can only operate after the  $(k-1)$ -th task completes its operation. Using start-up ( $z_{i,\tau}^k$ ) and shut-down ( $y_{i,\tau}^k$ ) indicators, we formulate these operation

dependence constraints as follows.

$$\sum_{\tau=0}^{N-1} \tau \cdot z_{i,\tau}^k \geq \sum_{\tau=0}^{N-1} \tau \cdot y_{i,\tau}^{k-1}, \quad \forall i, k = 2, 3, \dots, ST_i. \quad (2.8)$$

where  $\sum_{\tau=0}^{N-1} \tau \cdot z_{i,\tau}^k$  represents the duration before task  $k$  starts up and  $\sum_{\tau=0}^{N-1} \tau \cdot y_{i,\tau}^{k-1}$  represents the duration before task  $k - 1$  shuts down. Note that  $\tau$  denotes the relevant time slot index within the time horizon (e.g.,  $\tau = 0, 1, \dots, N - 1$ ), and the request of the event happens at beginning of the time horizon ( $\tau = 0$ ).

**(6) Delay constraints:** As shown in the dependence constraints (2.8), an interval between sequential tasks may be allowed. Different tasks have different delay interval tolerances depending on the nature of the task. For example, after the cycle of pumping the water to fill the drum of the clothes washer, it is beneficial to wait for a while to make full use of the cleaning effect of the detergent, but waiting too long is harmful to the clothes. The following task delay constraints guarantee that these task delay flexibilities will not impact the task quality.

$$\sum_{\tau=0}^{N-1} \tau \cdot z_{i,\tau}^k - \sum_{\tau=0}^{N-1} \tau \cdot y_{i,\tau}^{k-1} \leq D_i^k, \quad \forall k, \quad (2.9)$$

where  $D_i^k$  is the delay tolerance for task  $k$  of event  $i$ .

Besides each individual task delay tolerance, the customer usually specifies an overall event delay, i.e., once the customer pushes the start button of an appliance, how much time he/she is willing to wait for until event completion. This total delay constraint can be obtained by adding all delays of tasks for this event, as shown in (2.10).

$$\sum_{k=2}^{ST_i} \left( \sum_{\tau=0}^{N-1} \tau \cdot z_{i,\tau}^k - \sum_{\tau=0}^{N-1} \tau \cdot y_{i,\tau}^{k-1} \right) + \sum_{\tau=0}^{N-1} \tau \cdot z_{i,\tau}^1 \leq D_i, \quad \forall i, \quad (2.10)$$

where  $D_i$  is the overall delay tolerance of task  $i$ .

The optimization can be formulated as the objective function in (2.1) under constraints (2.2) – (2.10), which together form an integer linear programming (ILP) problem.

### 2.3.2 PEV battery charging/discharging scheduling

The PEV battery can either act as the demand – when in the charging mode – or the supply of electricity when in the discharging mode. The flexibility of the PEV batteries is not only from the charging/discharging status switching, but also from the power drawn to and from the battery. The scheduling time horizon for the PEV battery is from its arrival to the departure while it is plugged in. The BEMC aims to minimize the energy cost of charging during this period, while guaranteeing the battery being charged above some level at the departure time.

To capture the two possible modes of a PEV battery, we employ a binary variable  $v_\tau$  to indicate mode status at time slot  $\tau$ ; specifically,  $v_\tau = 0$  for charging mode and  $v_\tau = 1$  for discharging mode. Two ancillary variables  $sc_\tau$  and  $sd_\tau$  are defined to indicate start charging and start discharging actions. Let continuous variable  $B_\tau$  denote charging/discharging rate, where positive value is for charging and negative value is for discharging. Since there is an efficiency for both charging and discharging, we define an ancillary variable  $\tilde{B}_\tau$  to indicate power drawn to (charging, positive value) or from (discharging, negative value) the PEV battery. In the following parts we model operation constraints for battery and formulate the battery scheduling optimization to minimize the energy cost to customers.

**(1) Coupling status constraints:** The coupling constraints of battery status, start charging and start discharging indicator, which are similar to those of the appliances' ON/OFF status and start-up/shut-down, can be written as

$$v_\tau - v_{\tau-1} = sd_\tau - sc_\tau, \quad \forall \tau, \quad (2.11)$$

$$sd_\tau \leq v_\tau, \quad \forall \tau, \quad (2.12)$$

$$sc_\tau \leq (1 - v_\tau), \quad \forall \tau, \quad (2.13)$$

$$v_\tau, sc_\tau, sd_\tau \in \{0, 1\}, \quad \forall \tau. \quad (2.14)$$

**(2) Minimum charging/discharging time constraints:** To protect the lifespan of the PEV battery, the frequent charging/discharging switch should be avoided. We employ a minimum charging/discharging time denoted by  $CT/DT$ , which indicates if the battery starts charging/discharging mode, it must remain in the same mode for at least  $CT/DT$  time slots, respectively. These constraints are related to the battery status, the start charging and the start discharging indicators as follows:

$$\sum_{\tau'=\max(0,t-DT+1)}^{\tau} sd_{\tau'} \leq v_\tau, \quad \forall \tau \quad (2.15)$$

$$\sum_{\tau'=\max(0,t-CT+1)}^{\tau} sc_{\tau'} \leq 1 - v_\tau, \quad \forall \tau \quad (2.16)$$

**(3) Charging/discharging rate constraints:** The charging and discharging rate should be bounded by  $[B_c^l, B_c^u]$  and  $[B_d^l, B_d^u]$ , respectively. Note that for discharging,  $B_d^l, B_d^u \leq 0$ . The charging/discharging status  $v_\tau$  can be used to write

this constraint in a uniform way.

$$(1 - v_\tau)B_c^l + v_\tau B_d^l \leq B_\tau \leq (1 - v_\tau)B_c^u + v_\tau B_d^u, \quad \forall \tau. \quad (2.17)$$

We can see when  $v_\tau = 1$  (discharging),  $B_\tau \in [B_d^l, B_d^u]$ , and when  $v_\tau = 0$  (charging),  $B_\tau \in [B_c^l, B_c^u]$ .

**(4) State of charge (SOC) constraints:** Due to an efficiency parameter for both charging and discharging, the charging/discharging rates  $B_\tau$  are not the same as the power drawn to/from the battery  $\tilde{B}_\tau$ . Let  $\eta_I$  and  $\eta_O$  denote the charging and discharging efficiency, respectively. The relation of charging/discharging rates and power drawn to/from the battery in terms of the efficiency can be written as:

$$\tilde{B}_\tau = \begin{cases} \eta_I \cdot B_\tau, & \text{if } B_\tau \geq 0, \\ \frac{B_\tau}{\eta_O}, & \text{if } B_\tau < 0, \end{cases} \quad (2.18)$$

where the coefficients  $\eta_I$  and  $\eta_O$  lie in  $(0, 1]$ . However, (2.18) depends on the actual value of charging/discharging rates  $B_\tau$ , which are decision variables, so they cannot be added as constraints directly. In order to express these conditional constraints as linear constraints, we introduce big values ( $M_{B1} - M_{B4}$ ) as an ancillary constant, and convert (2.18) into the following four inequalities:

$$\tilde{B}_\tau \geq B_\tau \cdot \eta_I - v_\tau \cdot M_{B1}, \quad (2.19)$$

$$\tilde{B}_\tau \geq \frac{B_\tau}{\eta_O} - (1 - v_\tau) \cdot M_{B2}, \quad (2.20)$$

$$\tilde{B}_\tau \leq B_\tau \cdot \eta_I + v_\tau \cdot M_{B3}, \quad (2.21)$$

$$\tilde{B}_\tau \leq \frac{B_\tau}{\eta_O} + (1 - v_\tau) \cdot M_{B4}. \quad (2.22)$$

These ancillary variables should be set as small as possible for to accelerate the speed. According to the structures of these inequalities, we can set  $M_{B1} = -\frac{B_d^u}{\eta_O}$ ,  $M_{B2} = \frac{B_c^u}{\eta_O}$ , and  $M_{B3} = M_{B4} = 0$ . It can be easily verified that with the values of these ancillary variables, deterministic inequalities (2.19) – (2.22) are equivalent to the conditional equalities in (2.18).

Let  $EC$  denote the capacity of the PEV battery, and the state of charge (SOC)  $\zeta$  is defined as the battery energy level normalized by the battery capacity  $EC$ . Assume the SOC of the battery at time of arrival is  $\zeta_0$ . During charging/discharging over the time horizon  $L$ , lower and upper bounds of SOC,  $[\zeta^-, \zeta^+]$ , are specified to help protect the battery. For example, the lead-acid battery should not be discharged below 50% [27], and for the Lithium-based battery, deep discharge (depletion) should be avoided to prolong battery's lifespan [28].

$$\zeta^- \leq \frac{\zeta_0 \cdot EC + \sum_{\tau'=1}^{\tau} \tilde{B}_{\tau'}}{EC} \leq \zeta^+, \quad \tau = 1, \dots, L. \quad (2.23)$$

At the time of departure, i.e.,  $\tau = L$ , additional SOC constraint is needed such that the SOC of PEV battery needs to be above a certain level  $\zeta_d$  specified by the customer, i.e.,

$$\zeta_d \leq \frac{\zeta_0 \cdot EC + \sum_{\tau'=1}^L \tilde{B}_{\tau'}}{EC} \leq \zeta^+. \quad (2.24)$$

If electricity prices are positive, the left inequality in (2.24) is always binding to incur lowest cost. But for negative electricity prices, the battery may be charged more at departure to earn money, or to offset paying for the same charge during future intervals with positive prices.

**(5) Optimization formulation:** The BEMC schedules battery operation to minimize electricity cost, i.e.,

$$\min_{B_\tau} \sum_{\tau=1}^L B_\tau \cdot \kappa_\tau. \quad (2.25)$$

Note that when in discharging mode, the customer actually sells back electricity at whatever the price is at the time. The optimization for battery charging/discharging scheduling can be formulated given the objective (2.25) and constraints (2.11) – (2.24), which is a mixed-integer linear programming (MIP) problem.

## 2.4 Thermal Appliance Scheduling

Unlike washing machines or dishwashers, thermal appliances’ scheduling is related to the thermal dynamics of the building and the customer’s comfort level. In the following sections, we provide specifications of a building room prototype and model the thermal dynamics of the building room using a set of linear differential equations and the customer’s comfort level using the predictive mean vote (PMV) index. Then we study how these two factors are incorporated in the scheduling optimization for thermal appliances.

### 2.4.1 Thermal dynamics of building room

We consider a case from a multi-family home which has at least five housing units, each of which must share at least a floor or a ceiling with another unit. The envelope geometry of this room prototype is shown in Figure 2.3(a). We apply a

typical floor plan assuming a single zone for the entire unit and no heat transfer between the units, which is reasonable, since they are all similarly conditioned. The building floor-to-floor height is assumed to be three meters. All the other specifications of the building room prototype are based on the National Renewable Energy Laboratory (NREL) building America house simulation protocols [29], containing the physical property of walls, floor, windows, and the typical parameters for infiltration, internal gain, and system efficiency. These specifications can be used to compute the thermal mass capacity and heat transfer coefficient described in the next part. Due to the limit of space, these specifications are omitted here and the reader can refer to [29] for more details.

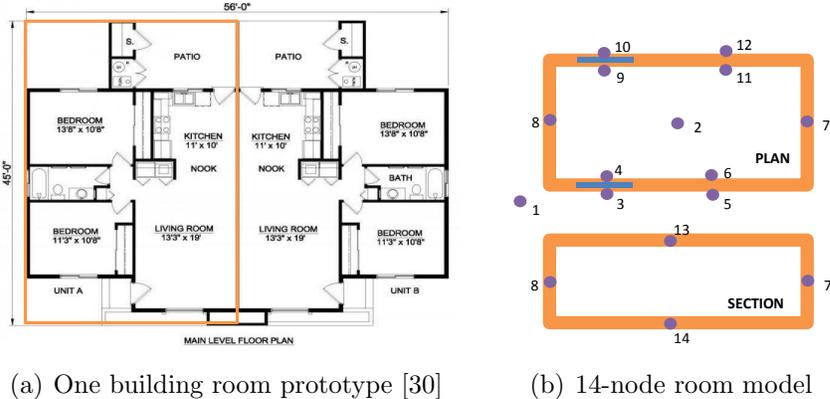


Figure 2.3: Building room prototype and model

When thermal appliances, such as the electric heater and A/C are scheduled by the BEMC, the whole building is a thermal dynamic system; the walls, floor, and ceiling transfer heat to the room while serving as a thermal mass for heat storage. The building thermal dynamic characteristics determine the heating and cooling demands of the room under specified comfort levels and given weather conditions. To optimize the schedules of the thermal appliances, we model the building thermal system using a finite element discretization technique. The technique discretizes the space into nodes and an element connecting nodes represents one heat transfer

phenomena [31]. Figure 2.3(b) shows a simple 14-node model of the prototype house; node 1 represents the ambient environment, which is used as the boundary conditions for the thermal dynamics. Hence, we regard the temperature for node 1 to be known and use the typical meteorological year (TMY) weather data to quantify it [32].

Except for node 1, the heat transfer equation per node can be formulated as

$$M_j \frac{dT_j(\tau)}{d\tau} + \sum_{k=1, k \neq j}^{14} S_{jk} [T_j(\tau) - T_k(\tau)] = f_j(\tau), \quad (2.26)$$

where  $M_j$  is the heat capacity of a building component represented by node  $j$  and  $M_j \frac{dT_j(\tau)}{dt}$  is the dynamic heat change due to the thermal inertia of the mass;  $S_{jk}$  is the heat transfer coefficient from node  $j$  to node  $k$ , by either conduction, convection, or radiation.  $T_j(\tau)$  is the temperature of node  $j$  at time slot  $\tau$ .  $f_j(\tau)$  is the external heat gain to node  $j$  at time slot  $\tau$ , i.e., solar radiation, heat dissipation from lighting and appliances, and the thermal appliances like the heater and A/C.

The thermal equations for all nodes at time slot  $\tau$  can be written in a matrix form as (2.27).

$$\mathbf{M} \frac{d\bar{T}(\tau)}{d\tau} + \mathbf{S} \bar{T}(\tau) = \bar{f}(\tau), \quad (2.27)$$

where  $\mathbf{M} = \text{diag}\{M_j\}$  is the thermal inertia matrix and  $\mathbf{S}$  is the heat transfer coefficient matrix. Vector  $\bar{T}(\tau)$  and  $\bar{f}(\tau)$  represent temperature and external heat gain for all nodes. Since the ambient environment (node 1) is used as the boundary conditions for the thermal dynamics system, we remove the first row of  $\mathbf{M}$  and  $\mathbf{S}$  so that only thermal dynamics of node 2 – 14 are formulated.

The external heat gain can be divided into thermal appliances, lighting, non-

thermal appliances, and solar radiation, i.e.,  $\bar{f} = \bar{f}_{\text{solar}} + \bar{f}_{\text{lighting}} + \bar{f}_{\text{app}} + \bar{f}_{\text{thermalapp}}$ . Heat gains from solar radiation and lighting are considered as known whereas heat gain from appliances is associated with decision variables in the optimization scheme.

## 2.4.2 PMV comfort level

Another important constraint is the human comfort level, which represents the thermal flexibility of the customer. We employ the PMV index [33]. PMV is an index that predicts the mean value of the votes of a large group of persons on the 7-point thermal sensation scale, i.e.,  $[-3,3]$  represents 7 thermal sensations from cold to hot, to describe the thermal comfort level.

The calculation of PMV can be referred to the ISO 7730 standard [33], which depends on several factors such as human metabolic rate, clothing insulation, clothing surface area factor, air temperature, mean radiant temperature (MRT), relative air velocity, and relative humidity. In our studies, we evaluate how the temperatures (air temperature  $t_a$  and (MRT)  $\bar{t}_r$ ) influence the PMV value, so we set all other factors to be constant nominal values as shown in Table 2.1.

Table 2.1: Parameters for PMV computation

Parameter	Value
Metabolic rate	1 met = 58.2 W/m <sup>2</sup>
Clothing insulation	1 clo = 0.155 m <sup>2</sup> ·°C/W
Relative air velocity	0.5 m/s
Relative humidity	40 %

However, the calculation of PMV in [33] requires iterative numerical methods to solve equations, which cannot be integrated into our optimization directly. Alter-

natively, we vary the  $t_a$  and  $\bar{t}_r$  from their normal range 10°C–30°C and 10°C–40°C, respectively, and compute PMV values using the iterative algorithm in [33], which is shown in Figure 2.4.

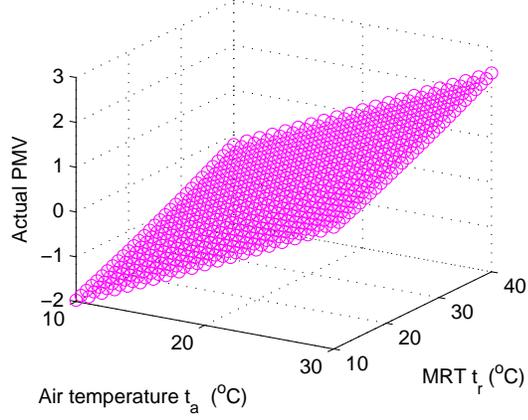


Figure 2.4: PMV data computed using iterative algorithm

We observe that all the PMV data points in Figure 2.4 follow a plane, so we apply linear regression to estimate the two-variable linear function of PMV in terms of  $t_a$  and  $\bar{t}_r$ , which gives the function as

$$\text{PMV} = -3.923 + 0.133 \cdot t_a + 0.0577 \cdot \bar{t}_r, \quad (2.28)$$

with an error variance as low as  $4.5113 \times 10^{-4}$ .

### 2.4.3 Optimization formulation

The BEMC aims to schedule the thermal appliances' power consumption  $Q_\tau$  to minimize energy cost, given that the PMV is within a customer specified range

over time horizon  $N$ , i.e.,

$$\min_{Q_\tau} \sum_{\tau=0}^{N-1} Q_\tau \cdot \kappa_\tau, \quad (2.29)$$

where  $Q_\tau$  needs to follow the specification of the electric heater or A/C as

$$Q^- \leq Q_\tau \leq Q^+, \quad \forall \tau, \quad (2.30)$$

where  $Q^-$  and  $Q^+$  are the lower and upper bounds of power consumption, respectively.<sup>1</sup>

Thermal dynamic equations in (2.27) serve as equality constraints on the room's thermal evolution. To convert them into linear constraints, we employ approximated difference equations

$$\mathbf{M}(\bar{T}_\tau - \bar{T}_{\tau-1}) + \mathbf{S}\bar{T}_\tau = \bar{f}_\tau, \quad \forall \tau. \quad (2.31)$$

The external heat gain of thermal appliance  $f_{\text{thermalapp},\tau}$  in  $f_\tau$  takes effect in room space (node 2), with the value of  $\beta Q_\tau$ , where  $\beta$  is the heat efficiency for the thermal appliance. So  $\bar{f}_{\text{thermalapp},\tau}$  can be expressed as  $(0, \beta Q_\tau, 0, \dots, 0)^T$ . If the thermal appliance is A/C for cooling, we can simply change the sign of  $\bar{f}_{\text{thermalapp},\tau}$ .

Given a PMV range  $[PMV^-, PMV^+]$  specified by the customer, the comfort level constraint can be formulated using the linear regression result in (2.28) for all time slots as

$$PMV^- \leq -3.923 + 0.133t_a + 0.0577\bar{t}_r \leq PMV^+, \quad \forall \tau, \quad (2.32)$$

---

<sup>1</sup>For the thermal appliances that only offer ON/OFF operation flexibility, binary variables can be introduced, and the optimization problem will change to a mixed-integer programming (MIP) problem.

where  $t_a$  is the room air temperature ( $T_2$ ), and the MRT  $\bar{t}_r$  of the room can be expressed in an empirical formula [34] as

$$\begin{aligned} \bar{t}_r = & [0.18 \times (T_{13}+T_{14})+0.22 \times (T_7+T_8)+(0.3-0.01) \\ & \times (T_6+T_{11})+0.01 \times (T_4+T_9)] / [2 \times (0.18+0.22+0.3)]. \end{aligned}$$

This PMV constraint in (2.32) is a linear inequality constraint. Combining the objective in (2.29) and constraints (2.30) – (2.32) gives a linear programming problem.

In the external heat gain vector  $\bar{f}$ , the solar radiation and the lighting part  $\bar{f}_{\text{solar}}$ ,  $\bar{f}_{\text{lighting}}$  can be viewed as fixed value inputs. The thermal appliance part  $\bar{f}_{\text{thermalapp}}$  is the decision variable. The non-thermal appliance external heat gain part  $\bar{f}_{\text{app}}$  can be divided into heat gain from non-schedulable (critical) appliances' operation,  $\bar{f}_{\text{app}}^n$ , and schedulable appliances' operation,  $\bar{f}_{\text{app}}^s$ . The latter depends on the scheduling optimization discussed in Section III, i.e., there are coupling effects on non-thermal appliances' scheduling and thermal appliances' scheduling. When the customer requests to turn on a non-thermal scheduleable appliance (event  $i$ ), the following objective function is formulated

$$\min_{u_{i,\tau}^k, y_{i,\tau}^k, z_{i,\tau}^k, Q_\tau} \sum_{\tau=0}^{N-1} \left( \sum_{k=1}^{ST_i} u_{i,\tau}^k \cdot P_i^k + Q_\tau \right) \cdot \kappa_\tau, \quad (2.33)$$

where the term in parenthesis denotes the total power consumption of both the schedulable non-thermal appliance and thermal appliance. Note that the time horizon  $N$  in Section III and Section IV may be different, so we set the time-horizon in (18) as the maximum of these two values. The constraints for (18) are (2) – (10) and (15) – (17) except that in constraint (16), the heat gain from non-thermal schedulable appliances (all tasks from event  $i$ ) is applied to the room air

(node 2), i.e.,

$$\bar{f}_{\text{app},\tau}^s = [0, \sum_{k \in ST_i} \gamma_i^k u_{i,\tau}^k P_i^k, 0, \dots, 0]^T, \quad (2.34)$$

where  $\gamma_i^k$  is the ratio of heat dissipation of task  $k$  of event  $i$  to its power consumption. Once this appliance has been scheduled, the corresponding heat dissipation part will be added to the heat gain from non-scheduleable appliances ( $\bar{f}_{\text{app}}^n$ ).

## 2.5 Numerical Results

In this section, we evaluate our proposed MPC-based appliance scheduling scheme. Since current time-varying retail market prices are based on wholesale electricity prices, in our simulations we employ the wholesale market prices directly as a proxy for the time-varying retail prices.<sup>2</sup> Specifically, we pick three pricing schemes for city of Chicago (PJM) [35], and New York City (NYISO) [36], [37] shown in Table 2.2. The means  $\bar{\kappa}$  and standard deviations  $\kappa_{\text{std}}$  of these three pricing schemes are also compared in Table 2.2. We can see pricing scheme 1 has a lower average

Table 2.2: Pricing schemes comparison

Pricing	Scheme 1	Scheme 2	Scheme 3
Source	Chicago [35]	N.Y.C [36]	N.Y.C [37]
Type	day-ahead (hourly)	day-ahead (hourly)	real-time (10-min)
$\bar{\kappa}$ (\$/MWh)	36.2324	75.8323	75.7839
$\kappa_{\text{std}}$	9.5389	32.7524	105.3502

price compared to pricing schemes 2 and 3. With the similar average price, pricing scheme 3 has a higher standard deviation than that of pricing scheme 2.

To explore the energy-cost savings potential of our proposed scheme, we deploy

<sup>2</sup>The price forecasts are assumed to be perfect as the actual values in the simulation; thus the results serve as upper bounds for the MPC method. Similar assumptions apply for the weather forecast.

the prototype residential case in two distinct climate zones [32]: Chicago and Dallas, whose mean temperatures  $\bar{T}$  in winter and summer are shown in Table 2.3. We can see that Chicago has a lower temperature in the winter, while Dallas has a higher temperature in the summer.

Table 2.3: Average temperature comparison for two places

Data	Season	Chicago	Dallas
$\bar{T}$	Winter	-3.97°C	7.18°C
	Summer	22.42°C	30.80°C

Washing machines, clothes dryers, and dishwashers are delay flexible appliances considered in the simulation. We categorize them into clothes washing event (Event 1) and dish-washing event (Event 2), each of which having four tasks. Note that both the washing machine and clothes dryer are included in the clothes washing event, where Task 4 is the operation of the clothes dryer. The parameters including operation time, power consumption, and delay tolerance of each task are shown in Table 2.4.<sup>3</sup> The customers' request times of these events are assumed to be uniformly distributed from 10:00 AM to 10:00 PM in the simulation; however, in practice, this distribution can vary according to customers' power usage patterns.

Table 2.4: Appliances (events) parameters

Event	Task	Operation time	Power	Delay tolerance
E 1	Task 1	30 min	400 W	180 min
	Task 2	20 min	600 W	30 min
	Task 3	30 min	500 W	60 min
	Task 4	50 min	2400 W	90 min
E 2	Task 1	30 min	1000 W	240 min
	Task 2	40 min	1200 W	60 min
	Task 3	30 min	800 W	60 min
	Task 4	30 min	1300 W	120 min

<sup>3</sup>Since different devices have different specifications and operation modes, these parameters are assumed values.

The Nissan Leaf electric vehicle [38] is considered as the PEV prototype in the simulation, which has a battery capacity of 24 kWh sufficient for driving up to 100 miles. Both charging and discharging have a maximum rate of 3kW and an efficiency of 90%. We assume the arrival time is uniformly distributed between 5:00 PM to 7:00 PM, and departure time is uniformly distributed between 6:00 AM to 8:00 AM.

The duration of one time slot for scheduling is set to  $\Delta = 10$  minutes in the simulation, but the actual duration can be adjusted to other values in practice. In the simulation, we assume the BEMC obtains real data for both electricity prices and weather (temperature and solar radiation) in order to conduct the optimization. In practice, the cost savings to customers will be lower due to forecasting errors. The CPLEX 12.4 optimization solver [39] is used for MIP and linear programming problems formulated in Sections III and IV. MATLAB 2009a is used to formulate the problem and link the CPLEX solver. The simulation environment is of Intel Duo Core 2.0GHz with 2GB memory. The computation time for thermal appliances scheduling optimization for each iteration (LP) is 0.03s, and the time for joint non-thermal and thermal appliances scheduling optimization for each iteration (MIP) is 0.27s.

### **(1) One-day scheduling results**

Figure 2.5 shows the one-day scheduling results of the electric heater, clothes washer/dryer, and dishwasher for pricing scheme 2. We can see during a low-price period, the heater consumes more power and the PMV increases, i.e., the building temperature increases (but it is still within the comfort level range) to store heat. During the high-price period, the heater consumes less or zero power,

making use of the stored heat in the building.

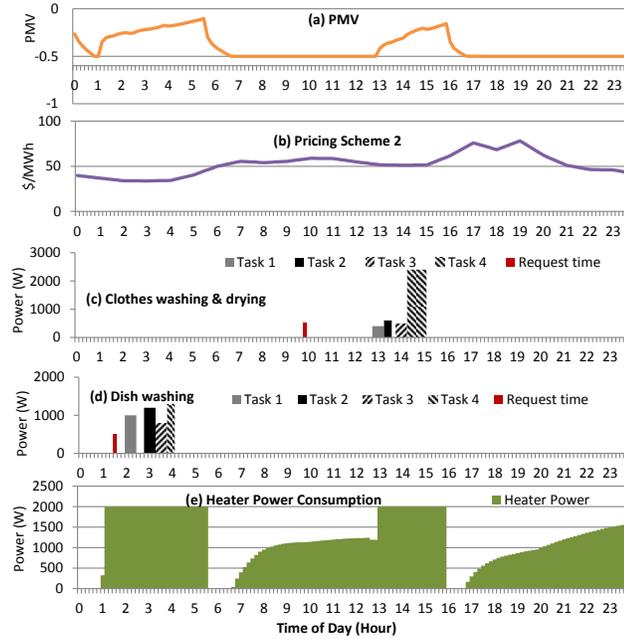


Figure 2.5: One-day scheduling results for pricing scheme 2

Figure 2.6 show battery charging/discharging scheduling results for pricing scheme 3. We can see that the optimization allows the battery to be charged at low price periods, and discharged at high price periods. By the time of departure, the customer is able to drive the PEV with proper SOC.

## (2) Cost saving to customers

We evaluate the monetary cost saving to a customer for our MPC-based appliance scheduling method over 40 consecutive days. Different pricing schemes and places with different temperature profiles are compared. To compare the energy cost saving, in the following sections, we also implement a benchmark scheduling scheme, in which non-thermal delay-flexible appliances are scheduled to operate without delay as customers request, and thermal appliances are scheduled to operate under

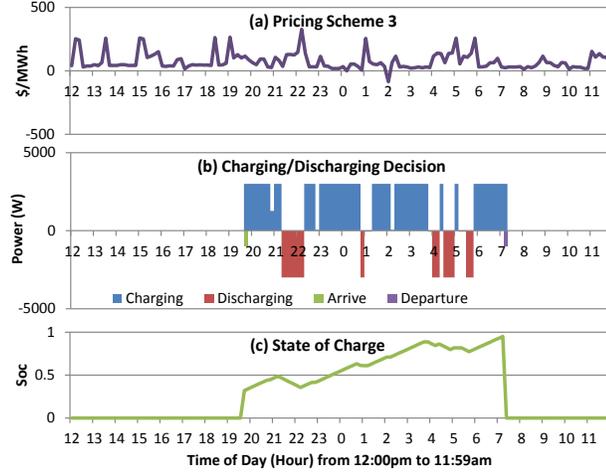


Figure 2.6: One-day battery scheduling result for pricing scheme 3

the same thermal dynamics and PMV comfort constraints without the objective of minimizing the energy cost.

Table 2.5 compares the energy cost using our proposed method for the three pricing schemes. We set the PMV comfort level bounds as  $[-0.5, 0.5]$ , and consider the heater in winter. We can see the electricity cost for pricing scheme 1 is much lower than that for scheme 2, due to the low values for both the mean and standard deviation of prices. With a similar mean value of prices for schemes 2 and 3, the higher standard deviation for scheme 3 results in a lower cost. The saving for scheme 3 almost reaches 20%.<sup>4</sup>

Table 2.5: Cost comparison for different pricing schemes

Pricing	Scheme 1	Scheme 2	Scheme 3
Energy Cost	\$34.09	\$74.96	\$56.42
Saving	10.76%	12.88%	19.74%

Figure 2.7 compares the energy cost (saving) for winter and summer cases of

<sup>4</sup>The actual saving will be less due to the assumption of the wholesale prices as the retail prices, and in reality the wholesale price is only part of the retail price. Also the accuracy of the model and prediction affect the cost saving.

Chicago and Dallas, using the same pricing scheme 1. We can see in the summer, since Dallas has a higher temperature (see Table 2.3), the energy cost is larger than that in Chicago. In the winter, however, energy cost in Chicago is larger due to the colder weather. In regard to the energy cost savings, which are also labeled in Figure 2.7, we see that these four cases give similar cost savings. Together with the results from Table 2.5, we can see the cost savings largely depend on the pricing schemes.

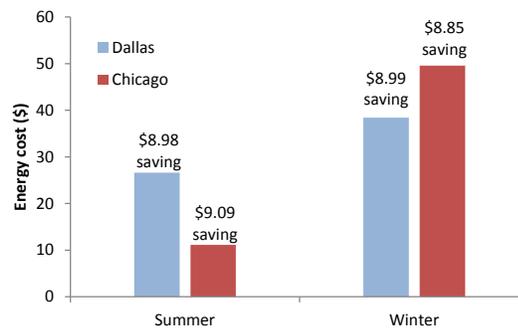


Figure 2.7: Energy cost comparison for winter and summer

### (3) Impact of customers' preference

Apart from the pricing and weather, the customers' preference over operational flexibilities of an appliance also plays an important role in energy cost savings. In this subsection, we evaluate the influence of thermal comfort (PMV) and delay flexibilities on energy cost savings.

Because customers will have diverse requirements in their comfort levels, the energy cost for the thermal appliance varies. For example, the stringent comfort range will not allow building thermal mass to store much energy;<sup>5</sup> thus, this provides less flexibility to hedge against price spikes or take advantage of lower or even

<sup>5</sup>Here the energy means heat in winter, or coolness in summer.

negative prices. Figure 2.8(a) shows the energy cost versus different PMV comfort levels for a 40-day simulation. We can see the more relaxed thermal comfort levels are, the less energy cost for customers. Consequently, these potential cost savings may also change the comfort preference for some customers.

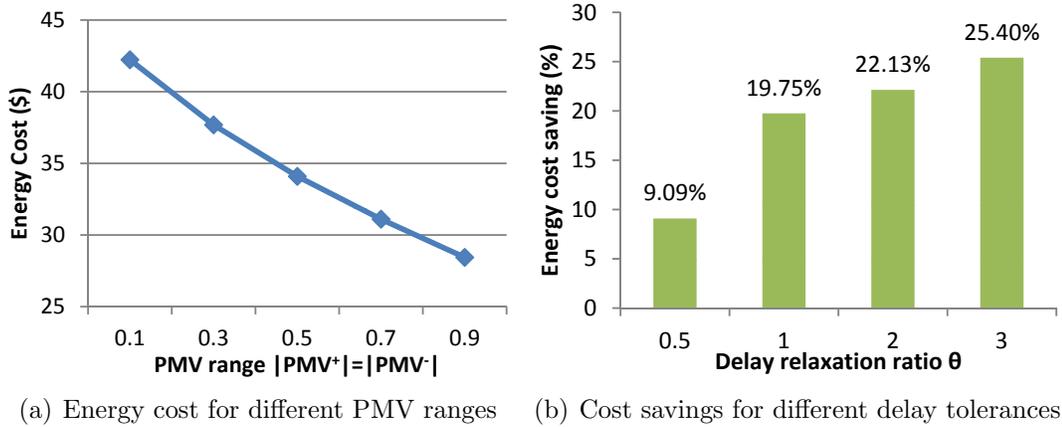


Figure 2.8: Impact of customers' preference

The customers may also have different delay tolerance flexibilities for their (non-thermal) appliances. Figure 2.8(b) shows how the delay flexibilities impact the energy cost savings. For a nominal delay tolerance  $D_i^k$  for each task, we vary a delay ratio  $\theta$  to uniformly describe the delay flexibilities, i.e., the delay tolerance is  $\theta \cdot D_i^k$ . We can see that the energy cost saving increases as delay tolerance is relaxed. In other words, customers with more (delay) operational flexibilities with their appliances will enjoy more energy savings. However, we can also see from Figure 2.8(b) that as delay tolerance increases, the cost savings saturate.

#### (4) Time horizon trade-off with price prediction uncertainty

In the MPC-based optimization, the time horizon  $N$  for planning is a significant parameter. On the one hand, employing longer time horizons will account for

more information in the future when making decisions; and thus, avoids short-sightedness. On the other hand, the future forecast information, i.e., electricity prices and weather data, usually have more uncertainty as the time horizon increases; since our model uses the deterministic forecast, the uncertainty in forecast incurs uncertainty to the results of our model. A trade-off lies between these two extremes for  $N$ .

In the simulation, we evaluate how the uncertainty of electricity prices affects this trade-off. Since the forecasts of electricity prices are not the focus of this paper, in the simulation we model the uncertainty of price forecasts as the actual data plus a noise distortion with a Gaussian distribution, i.e.,  $\kappa'_\tau = \kappa_\tau + \delta_\tau$ , where  $\delta_\tau \sim \mathcal{N}(0, \sigma_\tau^2)$ , and the variance  $\sigma_\tau$  increases with time  $\tau$ . The optimization is based on the distorted (forecast) electricity price data, but the billing applies actual prices. Figure 2.9 shows the energy cost over the 40 days versus the time horizon length,  $N$ , using pricing scheme 3.

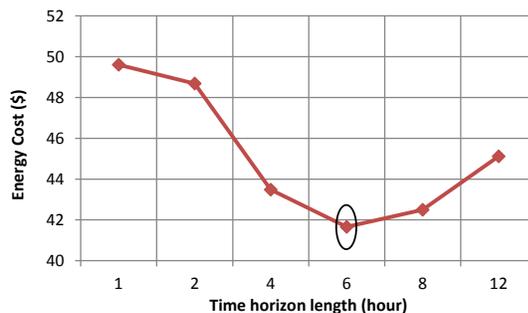


Figure 2.9: Time horizon length trade-off under price uncertainty

We can see when the time horizon length is small, the energy cost is higher due to not including much future information. As the time horizon length increases, the energy cost decreases. But excessively longer time horizons increase the energy cost as the uncertainty of electricity price forecasts increases. For Figure 2.9, the time horizon length of 6 hours gives the minimum energy cost. Note that Figure

2.9 is only an example to illustrate the time horizon length trade-off; however, the optimal time horizon length for the actual BEMC scheduling depends on the accuracy of the electricity price forecasts and variation of the prices.

## 2.6 Chapter Summary

In this chapter, an MPC-based appliance scheduling method is proposed for the residential BEMC assuming time-varying retail pricing in place. Several types of operational flexibilities of both non-thermal appliances and thermal appliances are exploited to help customers save on electricity bills. In our models, the operational dependence of non-thermal appliances is incorporated to further exploit price variations. In the thermal appliance optimization problem, the thermal mass of the building is modeled in thermal dynamics as linear difference equality constraints, in which the thermal appliances are scheduled smartly together with the thermal mass storage to incur potential savings from time-varying pricing. Simulation results demonstrate energy-cost savings using our method and the influence of customers' preferences is evaluated.

# Chapter 3

## Residential Appliance Scheduling: A Communication-based Control Approach

### 3.1 Introduction

In this chapter, we focus on how a consumer premise communication network (e.g., a home-area network) can be designed to permit load scheduling amongst flexible and controllable loads. In contrast to other existing models of EMCs, our proposed scheme alleviates the EMC of the communication and computation complexities of directly scheduling each individual load in the consumer premise. The functionalities of the EMC are reduced in our model to calculating a threshold maximum power consumption for the premise (e.g., the home). We cast the calculation of this time-varying threshold power as an optimization problem which accounts for price variation and uncertainty due to local wind generation. Our proposed joint

access and scheduling protocol describes how appliances access a common control channel for coordination so that this total maximum demand target is not exceeded for each time slot. Unlike existing local area network (LAN) media access protocols (e.g., Wi-Fi, G.hn, Zigbee), which could be used for one of several applications, our proposed scheme specifically addresses the load control problem. For in-home scenarios, relatively short control messages are occasionally exchanged; thus the approach can be easily implemented on higher data-rate LANs (such as those listed above) that support other applications as long as the load control messages are given transmission priority. To provide analytical foundations, we study the evolution of the protocol as a two-dimensional Markov chain and quantify the average delay experienced by individual appliances.

Besides improving demand side management, the role of the EMC also involves managing and making better use of local renewable distributed energy resources (DERs). Existing regulations incentivize residential and small business consumers to install renewable resources like solar panels and wind turbines [40]. Several products and installation services of residential renewable DER are also currently available [41]. As the penetration of these devices increases, the customers using them will become increasingly interested in how to best manage the local electricity supply and demand in the face of future pricing based DR mechanisms. To this end, we formulate the EMC's optimization of the target power value, which shows how this target value can be computed so that it accounts for both price variation and distributed wind power uncertainty. In this regard, we employ the Markov Chain Monte Carlo (MCMC) method to describe the stochastic property of power generation from wind turbines.

We note that unlike existing studies on EMC design (e.g., [9], [42], [43], [44]), the energy management approach presented here is tightly coupled with the underlying

communication algorithm used among the appliances. In this chapter, we present the whole picture of the proposed protocol, with detailed analysis, and we discuss how the protocol accommodates hard delay deadlines and any critical appliances that must be turned on without delay. Simulation results verify our analysis and demonstrate customers' savings using this proposed scheme.

The remainder of this chapter is organized as follows. Section 3.2 describes the joint appliance access and scheduling scheme for the EMC. Section 3.3 derives the average delay of an appliance using Markov chain model. Section 3.4 formulates an optimization for EMC to compute a target power budget. Numerical results are presented in Section 3.5, and we summarize our discussion in Section 3.6.

## 3.2 Access and Scheduling Scheme for Appliances

### 3.2.1 Protocol description

We assume a HAN with  $N$  appliances that share a common control channel. Time is segmented into *scheduling frames*, which are of the order of several (e.g., 5) minutes, and there may be multiple scheduling frames in the duration between target power updates. Let  $t$  denote the index of (scheduling) frame. Due to the long frame and short packet durations in the proposed scheme, we assume that collisions occur on the control channel with negligible probability.

We define the *active set* ( $\mathcal{A}_t$ ) as the set of appliances which are operating during frame  $t$ . We also assign a *leader* appliance ( $\ell_t$ ) that broadcasts beacon signals and makes admission control decisions during frame  $t$ . In a centralized scheme, the in-home EMC could be the (static) leader; or in a distributed scheme, any one of

the appliances in  $\mathcal{A}_t$  can be the leader. Later in this section, we present a possible leader assignment approach. Two types of appliances are considered in the home:

1. *Critical appliances*, that must join  $\mathcal{A}_t$  immediately. Examples include lighting, medical devices, laptops, TV, etc.
2. *Schedulable appliances*, that can be powered on with a tolerable delay.

We first consider the case with only schedulable appliances and then discuss modifications to include critical appliances. Figure 3.1 shows the frame structure of the proposed media-access and scheduling protocol. The frame consists of three phases: power update, power request and power scheduling, with duration  $T_1$ ,  $T_2$  and  $T_3$ , respectively.

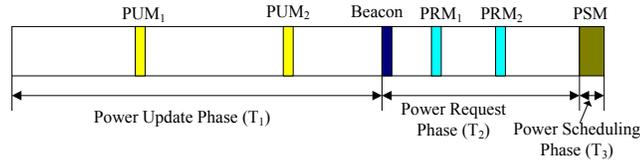


Figure 3.1: Frame structure of proposed protocol

**(1) Power update phase ( $T_1$ ):** Any active appliance which finishes its power usage sends a power update message (PUM) packet, indicating the amount of power the appliance is releasing. Any appliance that completes operation during the other two phases waits until the next power update phase to transmit the PUM packet. At the end of the power update phase, the leader appliance  $\ell_t$  calculates the remaining power budget:

$$\mathcal{P}_r^{(t)} = \mathcal{P}_{\max,t} - \sum_{j \in \mathcal{A}_t} \mathcal{P}_j, \quad (3.1)$$

where  $\mathcal{P}_r^{(t)}$  is the remaining power at time  $t$ ,  $\mathcal{P}_{\max,t}$  is the total demand target value and  $\mathcal{P}_j$  is the power usage for an appliance  $j$ . The leader then broadcasts this information in the beacon signal, which marks the start of the second phase.

**(2) Power request phase ( $T_2$ ):** This phase is divided into time slots. In the first slot, the leader appliance sends out a beacon signal indicating the remaining power budget  $\mathcal{P}_r^{(t)}$ . Any (schedulable) appliance,  $i$ , which wishes to join the active set listens to the control channel and, upon receiving the leader’s beacon signal, compares its requested power level with the remaining power budget. If its requested demand exceeds the remaining power budget, i.e.,  $\mathcal{P}_i \geq \mathcal{P}_r^{(t)}$ , then appliance  $i$  will defer requesting access to the active set until the next time frame. Otherwise, it attempts joining the active set in the current frame. Appliances that pass this “power check” enter into the *feasible appliance set* ( $\mathcal{F}_t$ ).

In order to be scheduled to power on, each appliance in  $\mathcal{F}_t$  waits for a random number of time slots (uniformly distributed over  $[1, w]$ , where  $w$  is the number of slots in this phase), and then sends out a power request message (PRM) packet to request the power usage. Appliances in  $\mathcal{F}_t$  whose PRM packets are received correctly by the leader enter into the *admissible appliance set* ( $\mathcal{D}_t$ ). During this time, the leader monitors the control channel and collects all the power request information.

**(3) Power scheduling phase ( $T_3$ ):** During this phase, the leader appliance makes an admission control decision, i.e., decides which (schedulable) appliances in  $\mathcal{D}_t$  can join the active set. These appliances form the *newly admitted appliance set* ( $\mathcal{N}_t$ ). Several admission control methods will be discussed later in this section. The leader then broadcasts a power scheduling message (PSM) packet, heard by all

(including active) appliances. Upon reception of the PSM packet, the appliances in  $\mathcal{N}_t$  start consuming power immediately. Any non-admitted appliances will wait until next frame to request again.

In the case that there is no EMC in the home (e.g., if the aggregator/utility determines the target power and send this to the home), the first appliance which powers on will be the leader and sends out the beacon signal to synchronize all appliances in the system. Subsequently, the last admitted appliance (the one listed last in the PSM message) is the leader for the upcoming frame. Thus, the role of the leader could potentially change from one frame to the next.

### 3.2.2 Admission control methods (ACM)

When the total request power is greater than the remaining power budget, i.e.,  $\sum_{i \in \mathcal{D}_t} \mathcal{P}_i \geq \mathcal{P}_{\max,t} - \sum_{j \in \mathcal{A}_t} \mathcal{P}_j$ , the leader appliance makes an admission decision. We consider the following ACMs, though others are possible:

**(1) Random ACM:** The leader appliance randomly removes appliances from the admissible set  $\mathcal{D}_t$  to meet  $\mathcal{P}_{\max,t}$ :

Step 1: Randomly remove appliance  $i$  from  $\mathcal{D}_t$ . Let  $\mathcal{D}_t = \mathcal{D}_t \setminus \{i\}$  and calculate  $\mathcal{P}_l = \sum_{i \in \mathcal{D}_t \cup \mathcal{A}_t} \mathcal{P}_i - \mathcal{P}_{\max,t}$ .

Step 2: If  $\mathcal{P}_l \leq 0$ , set  $\mathcal{N}_t = \mathcal{D}_t$  and stop; else go to Step 1.

**(2) Max-power oriented ACM:** The leader appliance maximizes the total additional power supported without exceeding the remaining power budget:

$$\begin{aligned}
& \max_{x_i} \quad \sum_{i \in \mathcal{D}_t} x_i \mathcal{P}_i \\
& \text{s.t.} \quad \sum_{i \in \mathcal{D}_t} x_i \mathcal{P}_i \leq \mathcal{P}_{\max,t} - \sum_{j \in \mathcal{A}_t} \mathcal{P}_j \\
& \quad \quad x_i \in \{0, 1\}.
\end{aligned} \tag{3.2}$$

This is a classical 0 – 1 knapsack problem, and can be solved using various existing approaches, e.g., [45]. The newly admitted set can be chosen as  $\mathcal{N}_t = \{i | x_i = 1, i \in \mathcal{D}_t\}$ .

**(3) Max-appliance oriented ACM:** The leader appliance selects in a way to support maximum number of appliances:

$$\begin{aligned}
& \max_{x_i} \quad \sum_{i \in \mathcal{D}_t} x_i \\
& \text{s.t.} \quad \sum_{i \in \mathcal{D}_t} x_i \mathcal{P}_i \leq \mathcal{P}_{\max,t} - \sum_{j \in \mathcal{A}_t} \mathcal{P}_j \\
& \quad \quad x_i \in \{0, 1\}.
\end{aligned} \tag{3.3}$$

The optimization problem in (3.3) can be solved by first sorting the requested power of individual appliances in an ascending order. Without loss of generality, we assume that  $\mathcal{P}_1 \leq \mathcal{P}_2 \leq \dots \leq \mathcal{P}_{|\mathcal{D}_t|}$ , where  $|\cdot|$  denotes the cardinality of a set.

Then,

$$\mathcal{N}_t = \left\{ \{1, \dots, \max\{i_S\}\} \mid \sum_{k=1}^{i_S} \mathcal{P}_k \leq \mathcal{P}_{\max,t} - \sum_{j \in \mathcal{A}_t} \mathcal{P}_j \right\}.$$

### 3.2.3 Critical appliance consideration

When a critical appliance has a power request, it starts operating immediately and sends out PRM packets over the control channel. If there is sufficient power budget to support this appliance, the leader simply needs to adjust the remaining power level for use in admission control. Since all active appliances have access to the control channel, they too can update their measurement of the remaining power to account for the critical appliance's power usage.

On the other hand, if there is not enough remaining power to accommodate the current usage, then the leader must employ a *curtailment control method* (CCM). Towards this end, the leader constructs the *curtailable appliance set* ( $\mathcal{C}_t$ ) from the active set at each time  $t$ . If  $\mathcal{C}_t \neq \emptyset$ , the leader selects a member of  $\mathcal{C}_t$  that needs to be removed from  $\mathcal{A}_t$ . Similar to ACM, the CCM adopted by the leader has three criteria: *random curtailment*; *min-power curtailment*; and *min-appliance curtailment*. The reader is referred to [46] for details, where we also discussed the case when  $\mathcal{C}_t = \emptyset$ , and modification of the protocol when adding deadlines to the appliances.

## 3.3 Markov Chain Analysis of the Protocol

In this section, we provide analysis to quantify the average delay performance of our proposed scheme for schedulable appliances only. We focus on schedulable appliances only as this assumption eases analysis. To do so, we model its stochastic behavior as a two-dimensional, discrete-time Markov chain and derive its stationary distribution. First, we define some additional appliance sets as follows: *Released*

*appliance set* ( $\mathcal{R}_t$ ), comprising appliances that complete usage cycle during time  $t$ ; *Joining appliance set* ( $\mathcal{J}_t$ ), consisting of appliances that wish to join  $\mathcal{A}_t$  during time  $t$ ; *Arrival appliance set* ( $\mathcal{I}_t$ ), containing appliances that request to power-on during time  $t$ .

### 3.3.1 Power usage statistic

We assume a finite number  $L$  of power levels for all appliances. Let  $\varepsilon_l$  denote the  $l$ -th power level, and without loss of generality we assume  $\varepsilon_l = l\varepsilon$ . Appliances' power draws are assumed to be random, with probability mass function (PMF) for appliance with power level  $\mathcal{P} = l\varepsilon$  as  $p_l \triangleq \Pr(\mathcal{P} = l\varepsilon)$ .

For a set of  $\omega$  appliances and each appliance with power usage  $\mathcal{P}_i$ , let  $\mathcal{P}_\Sigma = \sum_{i=1}^{\omega} \mathcal{P}_i$  denote the overall power usage of appliances. Then the PMF of  $\mathcal{P}_\Sigma$  can be written as

$$\phi_\Sigma(\omega, l_0) \triangleq \Pr(\mathcal{P}_\Sigma = l_0\varepsilon) = \sum_{\Omega} p_{l_1} \cdot p_{l_2} \cdots p_{l_\omega}, \quad (3.4)$$

where

$$\Omega = \left\{ (l_1, \dots, l_\omega) \in \mathbb{L}^\omega \mid \sum_{i=1}^{\omega} l_i = l_0, \mathbb{L} = \{1, \dots, L\} \right\},$$

$$l_0 \in \{\omega, \omega + 1, \dots, \omega L\}.$$

The cumulative distribution function (CDF) of  $\mathcal{P}_\Sigma$  is, therefore, given by

$$\Pr(\mathcal{P}_\Sigma \leq \bar{\mathcal{P}}) = \sum_{l_0=\omega}^{\lfloor \bar{\mathcal{P}}/\varepsilon \rfloor} \phi_\Sigma(\omega, l_0). \quad (3.5)$$

### 3.3.2 Two-dimensional discrete-time Markov chain model

Define the state  $S(t) = \{|\mathcal{A}_t|, |\mathcal{J}_t| \mid |\mathcal{A}_t| + |\mathcal{J}_t| \leq N\}$ , for frame  $t$ , indicating the number of active appliances and the number of appliances that wish to power on, respectively. The evolution of  $S(t)$  can be modeled as two-dimensional discrete-time Markov chain. From the structure of the protocol (see Figure 3.1), we have

$$\begin{aligned}\mathcal{A}_t &= \mathcal{A}_{t-1} \cup \mathcal{N}_{t-1} \setminus \mathcal{R}_t \Rightarrow |\mathcal{A}_t| = |\mathcal{A}_{t-1}| + |\mathcal{N}_{t-1}| - |\mathcal{R}_t|; \\ \mathcal{J}_t &= \mathcal{J}_{t-1} \cup \mathcal{I}_t \setminus \mathcal{N}_{t-1} \Rightarrow |\mathcal{J}_t| = |\mathcal{J}_{t-1}| - |\mathcal{N}_{t-1}| + |\mathcal{I}_t|; \\ \mathcal{N}_t &\subseteq \mathcal{D}_t \subseteq \mathcal{F}_t \subseteq \mathcal{J}_t \Rightarrow |\mathcal{N}_t| \leq |\mathcal{D}_t| \leq |\mathcal{F}_t| \leq |\mathcal{J}_t|.\end{aligned}$$

Let  $\Theta_{\mathcal{A}, \mathcal{J}}^{(t)}(i, j)$  indicate the event that  $S(t) = \{i, j\}$ , i.e., in frame  $t$ ,  $|\mathcal{A}_t| = i$ , and  $|\mathcal{J}_t| = j$ . The state transition probability  $\Pr(\Theta_{\mathcal{A}, \mathcal{J}}^{(t)}(i, j) \mid \Theta_{\mathcal{A}, \mathcal{J}}^{(t-1)}(n, m))$  for this two-dimensional Markov chain can be obtained as follows:

$$\begin{aligned}& \Pr(\Theta_{\mathcal{A}, \mathcal{J}}^{(t)}(i, j) \mid \Theta_{\mathcal{A}, \mathcal{J}}^{(t-1)}(n, m)) \\ &= \Pr(|\mathcal{N}_{t-1}| - |\mathcal{R}_t| = i - n, |\mathcal{N}_{t-1}| - |\mathcal{I}_t| = m - j \mid \Theta_{\mathcal{A}, \mathcal{J}}^{(t-1)}(n, m)) \\ &= \sum_{k=0}^{k_0} \Pr(|\mathcal{N}_{t-1}| - |\mathcal{R}_t| = i - n, |\mathcal{N}_{t-1}| - |\mathcal{I}_t| = m - j, |\mathcal{N}_{t-1}| = k \mid \Theta_{\mathcal{A}, \mathcal{J}}(n, m)) \\ &= \sum_{k=0}^{k_0} \Pr(|\mathcal{R}_t| = k + n - i \mid |\mathcal{N}_{t-1}| = k, \Theta_{\mathcal{A}, \mathcal{J}}^{(t-1)}(n, m)) \\ &\quad \times \Pr(|\mathcal{I}_t| = k + j - m \mid |\mathcal{N}_{t-1}| = k, \Theta_{\mathcal{A}, \mathcal{J}}^{(t-1)}(n, m)) \times \Pr(|\mathcal{N}_{t-1}| = k \mid \Theta_{\mathcal{A}, \mathcal{J}}^{(t-1)}(n, m)) \\ &= \sum_{k=0}^{k_0} \frac{(\mu' T_f)^{k+n-i} (\lambda' T_f)^{k+j-m}}{(k+n-i)!(k+j-m)!} e^{-(\mu'+\lambda')T_f} \times \Pr(|\mathcal{N}_{t-1}| = k \mid \Theta_{\mathcal{A}, \mathcal{J}}^{(t-1)}(n, m)), \quad (3.6)\end{aligned}$$

where  $k_0 = \min\{m, N - n\}$ . Here, we assume each appliance's power request and power release to independently follow the Poisson distribution with rates  $\lambda$  and  $\mu$ , respectively. Then, given the number of active appliances in the previous frame

( $|\mathcal{A}_{t-1}| = n$ ) and the number of newly requested appliances in the previous frame ( $|\mathcal{J}_{t-1}| = m$ ), we see that  $|\mathcal{I}_t|$  and  $|\mathcal{R}_t|$  follow a finite-sized queuing system with rates  $\lambda' = (N - n - m)\lambda$  and  $\mu' = n\mu$ , respectively.

The second term in (3.6) can be written as

$$\begin{aligned}
& \Pr(|\mathcal{N}_t| = k | \Theta_{\mathcal{A}, \mathcal{J}}^{(t)}(n, m)) \\
&= \sum_{p=k}^m \sum_{q=k}^p \Pr(|\mathcal{N}_t| = k, |\mathcal{D}_t| = q, |\mathcal{F}_t| = p | \Theta_{\mathcal{A}, \mathcal{J}}^{(t)}(n, m)) \\
&= \sum_{p=k}^m \sum_{q=k}^p \underbrace{\Pr(|\mathcal{F}_t| = p | \Theta_{\mathcal{A}, \mathcal{J}}^{(t)}(n, m))}_{P_0} \times \underbrace{\Pr(|\mathcal{D}_t| = q | |\mathcal{F}_t| = p, \Theta_{\mathcal{A}, \mathcal{J}}^{(t)}(n, m))}_{P_1} \\
&\quad \times \underbrace{\Pr(|\mathcal{N}_t| = k | |\mathcal{D}_t| = q, |\mathcal{F}_t| = p, \Theta_{\mathcal{A}, \mathcal{J}}^{(t)}(n, m))}_{P_2}. \tag{3.7}
\end{aligned}$$

**(1) Calculation of  $P_0$ :** Given  $n$  active appliances and  $m$  appliances which wish to join  $\mathcal{A}_t$ ,  $P_0$  is the probability that  $m - p$  appliances can not pass the power check, i.e., their individual power request levels are greater than the remaining power budget. The probability that an appliance  $i \in \mathcal{J}_t$  has its individual power request levels less than the remaining power budget is given by

$$\alpha_n \triangleq \Pr\left(\mathcal{P}_i \leq \mathcal{P}_{\max, t} - \sum_{j \in \mathcal{A}_t} \mathcal{P}_j \mid |\mathcal{A}_t| = n\right) = \sum_{l_0=n+1}^{\lfloor \mathcal{P}_{\max, t}/\varepsilon \rfloor} \phi_{\Sigma}(n+1, l_0).$$

Thus,  $P_0 = \binom{m}{p} \alpha_n^p (1 - \alpha_n)^{m-p}$ .

**(2) Calculation of  $P_1$ :** Given  $p$  appliances in *feasible appliance set* ( $\mathcal{F}_t$ ),  $P_1$  is the probability that exactly  $q$  appliances' packets can be successfully received and  $p - q$  appliances collide with each other when transmitting the PRM packets during

the power request phase.

$$\begin{aligned}
P_1 &= \Pr(|\mathcal{D}_t| = q | |\mathcal{F}_t| = p) \\
&= \underbrace{\Pr(p - q \text{ collisions} | q \text{ successes}, |\mathcal{F}_t| = p)}_{P_{10}} \times \underbrace{\Pr(q \text{ successes} | |\mathcal{F}_t| = p)}_{P_{11}}, \quad (3.8)
\end{aligned}$$

where  $P_{11} = \frac{w!}{w^q(w-q)!}$ . Let  $h$  denote the number of successful power request packets sent among the remaining  $p - q$  appliances given the conditions in  $P_{10}$ . Therefore,  $P_{10} = \Pr(h = 0) = 1 - \Pr(h \geq 1)$ , where  $h \geq 1$  signifies that, among  $p - q$  appliances' packets, at least one packet can be successfully received. Further, these packets can only choose a back-off value among  $w - q$  numbers. Let  $G_i$  denote the event that the  $i$ -th appliance out of  $p - q$  successfully transmits its packet given the conditions in  $P_{10}$ . Thus,

$$\begin{aligned}
P_{10} &= 1 - \Pr(h \geq 1) = 1 - \Pr\left(\bigcup_{i=1}^{p-q} G_i\right) \\
&= 1 - \sum_{j=1}^{p-q} (-1)^{j-1} \binom{p-q}{j} \frac{(w-q)!(w-q-j)^{p-q-j}}{(w-q)^{p-q}(w-q-j)!}. \quad (3.9)
\end{aligned}$$

**(3) Calculation of  $P_2$**  Given  $n$  active appliances and  $q$  appliances in  $\mathcal{D}_t$ ,  $P_2$  is the probability that  $k$  appliances are granted admission by the leader appliance. We employ random ACM to analyze the admission success probability, where, given  $|\mathcal{D}_t| = q$ ,  $q$  steps may exist for admission control. The *admissible set* at the  $(i + 1)$ -th step is denoted  $\mathcal{D}_t^{(i)}$ . Initially, we have  $\mathcal{D}_t^{(0)} = \mathcal{D}_t$ . Figure 3.2 shows the

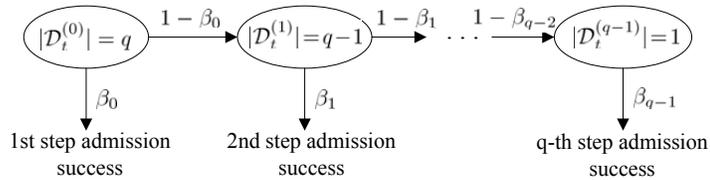


Figure 3.2: Admission success probability given  $|\mathcal{D}_t| = q$  and  $|\mathcal{A}_t| = n$

admission success probability given  $|\mathcal{D}_t| = q$  and  $|\mathcal{A}_t| = n$ .  $P_2$  can be regarded as the admission success probability at the  $(q - k + 1)$ -th step, and can be written as

$$\begin{aligned}
P_2 &= \Pr(|\mathcal{N}_t| = k | |\mathcal{D}_t| = q, |\mathcal{A}_t| = n) \\
&= \Pr(\text{admission success} | \Psi_{\mathcal{A},\mathcal{D}}(n, q), |\mathcal{D}_t^{(q-k)}| = k) \times \Pr(|\mathcal{D}_t^{(q-k)}| = k | \Psi_{\mathcal{A},\mathcal{D}}(n, q)) \\
&= \beta_{q-k} \prod_{j=0}^{q-k-1} (1 - \beta_j) \tag{3.10}
\end{aligned}$$

where  $\Psi_{\mathcal{A},\mathcal{D}}(n, q) = \{|\mathcal{A}_t| = n, |\mathcal{D}_t| = q\}$ . Given  $\mathcal{D}_t^{(i)}$ , the admission success probability  $\beta_i$  indicates the probability that the total request power level in  $\mathcal{D}_t^{(i)}$  is less than the remaining power budget, i.e.,

$$\beta_i \stackrel{\text{def}}{=} \Pr(\text{admission success} | |\mathcal{A}_t| = n, |\mathcal{D}_t^{(i)}| = q - i) = \sum_{l_0=q-i+n}^{\lfloor \mathcal{P}_{\max}/\varepsilon \rfloor} \phi_{\Sigma}(q - i + n, l_0), \tag{3.11}$$

$i = 0, \dots, q - 2$  and  $\beta_{q-1} = 1$ . Finally, given  $P_0$ ,  $P_1$  and  $P_2$ , we compute the state-transition probability given by (3.6).

### 3.3.3 Steady-state distribution

Let  $\pi_{n,m} \triangleq \lim_{t \rightarrow \infty} \Theta_{\mathcal{A},\mathcal{J}}^{(t)}(n, m)$  denote the steady-state distribution. For  $(N + 1)(N + 2)/2$  states, the steady-state distribution vector is  $\boldsymbol{\pi} \triangleq [\pi_{0,0}, \pi_{0,1}, \dots, \pi_{N,0}]$ . The steady-state distribution is obtained by solving  $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$  and  $\boldsymbol{\pi}\mathbf{1}^T = 1$ , where the state-transition probability matrix  $\mathbf{P}$  is given by (3.6). The marginal PMF for  $|\mathcal{A}_t|$  and  $|\mathcal{J}_t|$  can be computed as:  $\pi_n \triangleq \Pr(|\mathcal{A}_t| = n) = \sum_{m=0}^{N-n} \pi_{n,m}$ ; and  $\pi_m \triangleq \Pr(|\mathcal{J}_t| = m) = \sum_{n=0}^{N-m} \pi_{n,m}$ . For further analysis, we define the following

probabilities:

$$\xi_p \triangleq \Pr(|\mathcal{F}_t| = p) = \sum_{n=0}^N \sum_{m=0}^{N-n} \pi_{n,m} P_0, \quad (3.12)$$

$$\zeta_{n,q} \triangleq \Pr(|\mathcal{A}_t| = n, |\mathcal{D}_t| = q) = \sum_{m=0}^{N-n} \sum_{p=q}^m \pi_{n,m} P_0 P_1. \quad (3.13)$$

### 3.3.4 Delay analysis

Figure 3.3 shows the state transition diagram of the protocol, for one arbitrary appliance  $i$  in the system. Six states are considered for the appliance in this protocol.

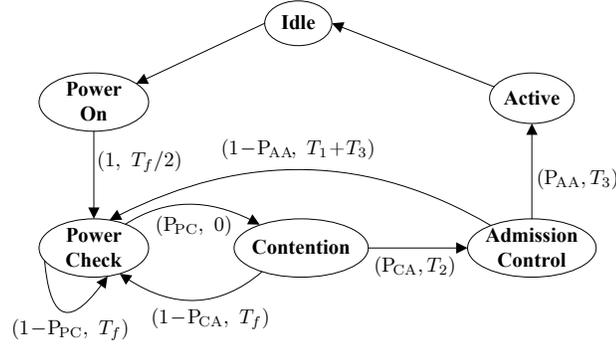


Figure 3.3: State transition diagram for an arbitrary appliance

An appliance is in the **Idle** state until it is requested into the **Power On** state by the customer. When an appliance is requested to power on, it first enters the **Power Check** state and correspondingly joins the *feasible appliance set* ( $\mathcal{F}_t$ ) in phase  $T_2$ . Next, the **Contention** state corresponds to joining *admissible appliance set* ( $\mathcal{D}_t$ ) in phase  $T_2$  (or returning to Power Check state if contention is unsuccessful). **Admission Control** is corresponds to joining *newly admitted appliance set* ( $\mathcal{N}_t$ ) in  $T_3$  phase and becoming an **Active** appliance (or returning to Power Check if admission is not granted). We now compute the power check success probability ( $P_{PC}$ ), contention success probability ( $P_{CA}$ ) and admission success probability ( $P_{AA}$ ).

(1)  $P_{PC}$ : Power check success indicates that an appliance  $i$  perceives its power request level less than the remaining power budget, i.e., this appliance is in the *feasible appliance set*  $\mathcal{F}_t$ . Then, the power check success probability  $P_{PC} \triangleq \Pr(i \in \mathcal{F}_t)$  is given by

$$P_{PC} = \sum_{n=0}^N \Pr(|\mathcal{A}_t| = n) \cdot \Pr(a \in \mathcal{F}_t | |\mathcal{A}_t| = n) = \sum_{n=0}^N \pi_n \alpha_n. \quad (3.14)$$

(2)  $P_{CA}$ : Contention success indicates that no collisions happen when an appliance  $i \in \mathcal{F}_t$  sends out the PRM packet and the *leader appliance* can receive it without errors, i.e., this appliance is in the *admissible appliance set*  $\mathcal{D}_t$ . From (3.12),  $P_{CA} \triangleq \Pr(i \in \mathcal{D}_t)$  is given by

$$\begin{aligned} P_{CA} &= \sum_{p=0}^N \Pr(|\mathcal{F}_t| = p) \times \Pr(\text{no packet collision} | |\mathcal{F}_t| = p) \\ &= \sum_{p=0}^N \xi_p \left(1 - \frac{1}{w}\right)^{[p-1]^+}, \end{aligned} \quad (3.15)$$

where  $[x]^+ = \max[0, x]$ .

(3)  $P_{AA}$ : Admission success indicates that the appliance  $i \in \mathcal{D}_t$  can be granted admission by the *leader appliance*, i.e., this appliance is in the *new active appliance set*  $\mathcal{N}_t$ . From (3.13),  $P_{AA} \triangleq \Pr(i \in \mathcal{N}_t)$  is given by

$$\begin{aligned} P_{AA} &= \sum_{n=0}^N \sum_{q=0}^{N-n} \Pr(|\mathcal{A}_t| = n, |\mathcal{D}_t| = q) \times \Pr(a \in \mathcal{N}_t | |\mathcal{A}_t| = n, |\mathcal{D}_t| = q) \\ &= \sum_{n=0}^N \sum_{q=0}^{N-n} \zeta_{n,q} \left[ \beta_0 + \sum_{i=1}^{q-1} \frac{q-i}{q-i+1} \beta_i \prod_{j=0}^{i-1} (1-\beta_j) \right]. \end{aligned} \quad (3.16)$$

We denote the average delay in the Power Check, Contention and Admission Control states as  $T_P$ ,  $T_C$  and  $T_A$ , respectively. From the state-transition diagram (see Figure 3.3), the average delay  $T_D$  is given by  $T_D = T_f/2 + T_P$ , where

$$T_P = P_{PC}T_C + (1 - P_{PC})(T_P + T_f), \quad (3.17)$$

$$T_C = P_{CA}(T_2 + T_A) + (1 - P_{CA})(T_P + T_f), \quad (3.18)$$

$$T_A = P_{AA}T_3 + (1 - P_{AA})(T_P + T_1 + T_3). \quad (3.19)$$

From (3.17) - (3.19), we have

$$T_D = \frac{T_f}{2} + \frac{T_f - P_{PC}P_{CA}P_{AA}T_1}{P_{PC}P_{CA}P_{AA}}. \quad (3.20)$$

We can see if all three probabilities are 1,  $T_D = \frac{3T_f}{2} - T_1$ , which is the overhead of a schedulable appliance joining the *Active Set* for this protocol. The smaller the product of these three probabilities  $P_{PC}P_{CA}P_{AA}$  is, the longer average delay an appliance will have.

### 3.4 Optimization of $\mathcal{P}_{\max,t}$ for EMC

In this section, we formulate an optimization problem that helps the EMC decide  $\mathcal{P}_{\max,t}$  integrating the distributed renewable generation.<sup>1</sup> The EMC aims to minimize the total cost of electricity on a daily basis while considering variations in electricity prices and uncertainty of local wind power generation. We assume that the EMC has hourly day-ahead prices for the entire day and receives real-time prices several minutes before each hour.

---

<sup>1</sup>We focus on the wind power here in the optimization; solar power can also be integrated if proper prediction algorithms can be applied.

### 3.4.1 Wind power statistics

We use wind speed data from the National Renewable Energy Laboratory (Solar Radiation Research wing), which is located at latitude  $39.742^\circ$  North and longitude  $105.18^\circ$  West. The elevation is 1828.8 meters. The data are gathered during the following period: January 1st, 2009 to December 31st, 2009, with 10 minutes sampling interval [47].

We model the stochastic behavior of wind power using the Markov Chain Monte Carlo (MCMC) technique, where the transition probability matrix is constructed from the historical wind speed data and wind turbine power curve. This matrix can be used to predict the wind power generation and aid in EMC scheduling framework. Given the wind speed, the power output of the wind turbine can be obtained from the power curves, which map the wind speed to wind turbine generation. For our model, we use the RAUM ENERGY<sup>TM</sup>3.5kW wind turbine, and the instantaneous system power curve is shown in its data-sheet [48], where the wind power outputs form a set  $\mathcal{H}$  with number of  $|\mathcal{H}| = 11$  states, i.e.,  $s_1, \dots, s_{11}$  representing the  $|\mathcal{H}|$  wind turbine generation values. For the first-order Markov Chain (FO-MC), the state evolution can be described by the transition probability matrix  $\mathbf{Q}$ , whose entry  $q_{ij}$  represents the transition probability from state  $i$  at frame  $t - 1$  to state  $j$  at frame  $t$ , i.e.,  $q_{ij} = \Pr\{X_t = j | X_{t-1} = i\}$ , which is independent of frame  $t$ . Second and third order Markov chains can also be used to describe the stochastic behavior. However, following the results in [49], these models provide approximately the same level of modeling accuracy as FO-MC. Thus, for simplicity in exposition, we choose FO-MC. Given the wind power data,

$q_{ij}$  can be estimated using the Monte Carlo method:

$$q_{ij} = \frac{n_{ij}}{\sum_{k \in \mathcal{H}} n_{ik}}, \quad (3.21)$$

where  $n_{ij}$  is the number of transitions from  $i$  to  $j$  encountered in the records. Figure 3.4 shows the transition probability matrix  $\mathbf{Q}$  obtained from the data. We notice that the diagonal and near-diagonal entries dominate, demonstrating the high auto-correlation of the wind power sequence.

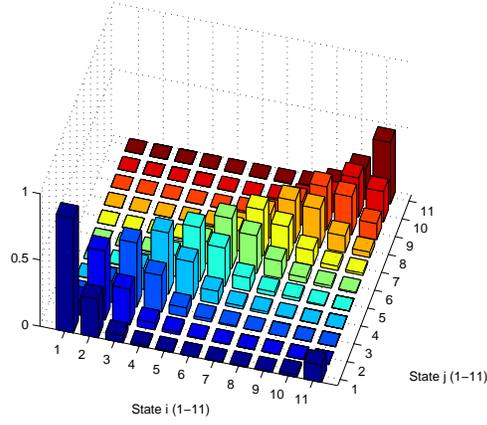


Figure 3.4: Transition probability matrix of wind power

### 3.4.2 Optimization formulation

Given  $\mathbf{Q}$  and the initial state  $s_{i0}$ , an estimate of the expected value of wind power at frame  $t$  is given by

$$\tilde{W}_t = \sum_{j=1}^{|\mathcal{H}|} s_j [\mathbf{Q}^t]_{i0,j}, \quad (3.22)$$

where  $[\mathbf{X}]_{i,j}$  indicates the entry with row  $i$  and column  $j$  in matrix  $\mathbf{X}$ . We formulate the optimization problem for the EMC at frame  $t_0$  with  $T$ -hour look ahead,  $K$ -min

time slots in a rolling-based manner. Note that  $T = 24$  is for daily optimization. For each frame  $t$ , the EMC optimizes  $\mathcal{P}_{\max,t}$  for  $Z = 60T/K$  cycles from  $t_0$ , i.e.,  $t \in [t_0 + 1, t_0 + Z]$ , with the predicted values of wind power  $\tilde{W}_{t_0+1} \dots \tilde{W}_{t_0+Z}$  using the actual wind power  $W_{t_0}$  as the initial state in (3.22). The problem can, therefore, be formulated as follows:

$$\begin{aligned}
& \min_{\mathcal{P}_{\max,t}} && \sum_{t=t_0+1}^{t_0+Z} \kappa_t (\mathcal{P}_{\max,t} - \tilde{W}_t) \\
& \text{s.t.} && \mathcal{L}_t + \tilde{W}_t \leq \mathcal{P}_{\max,t} \leq \mathcal{U}_t + \tilde{W}_t, \\
& && \sum_{t=t_0+1}^{t_0+Z} \mathcal{P}_{\max,t} \cdot \frac{K}{60} \geq \mathcal{E}_{t_0}.
\end{aligned} \tag{3.23}$$

Here,  $\kappa_t$  is the electricity price with  $K$ -minute the interval. Since wind power data intervals and pricing signal updates are 10-minute and 1-hour, respectively, the optimization interval  $K$  can be chosen as 10 minutes. The price  $\kappa_t$  during the first hour after  $t_0$  can be the real-time price (obtained several minutes before that hour). The following prices can either use the forecast hourly price as shown in [42], or the day-ahead price for approximation.  $\mathcal{L}_t$  and  $\mathcal{U}_t$  denote the lower and upper bounds, respectively, for power consumption level at frame  $t$ .  $\mathcal{E}_{t_0}$  denotes the minimum total power consumption for  $T$  hours horizon starting from  $t_0$ . These parameters can be determined by the EMC using learning algorithms that follow the consumer's usage patterns. Alternatively, we may select  $\mathcal{L}_t$  as the aggregate power consumption level by real-time appliances, i.e.,  $\mathcal{L}_t = \sum_{i \in \mathcal{V}_t} \mathcal{P}_i + \delta_t$ , where  $\mathcal{V}_t$  denotes the set of critical appliances at frame  $t$ , while  $\delta_t$  denotes the slack power level that corresponds to critical appliances with cycle power (e.g. air conditioners). Such a selection of  $\mathcal{L}_t$  guarantees sufficient power budget to accommodate real-time appliances. Additionally,  $\mathcal{U}_t$  may be provided by the utility/aggregator to avoid aggregate peak power. We note that, with additional (predicted) wind power, the

lower and upper bounds for  $\mathcal{P}_{\max,t}$  are adjusted accordingly in (3.23).

### 3.5 Numerical Results

We assume the time frame length  $T_f = 5$  min, where  $T_1 = 4.5$  minutes,  $T_2 = 29.9$ s and  $T_3 = 0.1$ s. The beacon signal and each packet are 0.1s long, and thus the number of slots in the power request phase is  $w = 298$  (not including the beacon signal). We consider  $L$  appliance types, and assume the power  $\mathcal{P}_i = k\varepsilon$ ;  $i = 1, \dots, N$ ,  $k = 1, \dots, L$ , which is uniformly distributed, i.e.,  $p_k \triangleq \Pr(\mathcal{P}_i = k\varepsilon) = 1/L$ , to match our assumption in Section 3.3. Note that the proposed protocol and our simulation approach itself do not require this uniform distribution; the assumption is made here to validate our analysis. In the simulation results presented, we choose  $L = 5$  and  $\varepsilon = 100$ W. For simplicity, we assume all  $L$  appliance types to have a duration of use that is exponentially distributed with mean  $1/\mu$ .

#### (1) Delay Characteristics

We assume  $\mathcal{P}_{\max,t} = \mathcal{P}_{\max}$  and a fixed arrival rate of  $\lambda$  for all hours. First, in Figure 3.5, we plot the average delay as a function of  $\mathcal{P}_{\max}$  for different  $(\lambda, \mu)$  combinations.

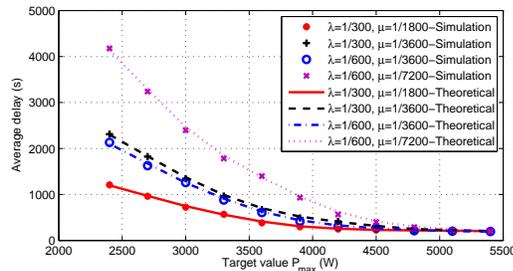


Figure 3.5: Average delay versus  $\mathcal{P}_{\max}$  with various  $(\lambda, \mu)$  (random ACM,  $N = 15$ )

Results demonstrate that the average delay (*i*) decreases with increasing  $\mathcal{P}_{\max}$ ; (*ii*) increases with appliance request (arrival) rate  $\lambda$ ; and (*iii*) decreases with power usage completion (departure) rate  $\mu$ . We also plot the theoretical average delay, given by (3.20), and see that it matches with the simulation results. We also note that, as  $\mathcal{P}_{\max}$  increases beyond a certain value, i.e., when the total power consumption without scheduling is at most  $\mathcal{P}_{\max}$ , the average delay converges to the overhead of the protocol  $\approx T_f/2 + T_2 + T_3 = 180s$ . This can also be seen from (3.20) corresponding to  $P_{PC}P_{CA}P_{AA} = 1$ .<sup>2</sup>

In Figure 3.6, we plot the average delay as a function of the number  $N$  of appliances with different ACMs. We see that the average delay increases as  $N$  increases, since larger number of appliances cause more congestion. We also note that the average delays of the three ACMs are comparatively close to each other. From the plot we conclude that, since random ACM has linear implementation complexity, it can be considered to be a desirable option.

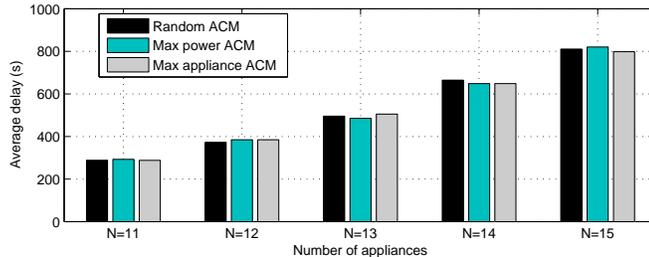


Figure 3.6: Average delay versus  $N$  with various ACMs. ( $\mathcal{P}_{\max} = 3000W$ ,  $\lambda = 1/900$ , and  $\mu = 1/3000$ )

Figure 3.7(a) shows the average delay versus the target power budget  $\mathcal{P}_{\max}$  with different number of critical appliances in the system. We only present results for the random CCM policy as the performance of other schemes matches closely with this. We see that with more critical appliances, the average delay for schedulable

<sup>2</sup>Since the number  $N$  of appliances is not large, the PRM packet receiving probability  $P_{CA} \approx 1$ .

appliances become larger. This is primarily due to the fact that when one schedulable appliance is curtailed by the EMC/leader appliance, it needs to re-compete with other appliances to join the active set. When accounting for all appliances, it is interesting to note that the average delay is almost the same as the case of no critical appliance, indicating the presence of critical appliances will not increase the average delay in the system. Figure 3.7(b) shows the average curtailment with different number of critical appliances. As expected, we see that with a larger number of critical appliances or smaller  $\mathcal{P}_{\max}$ , the average number of curtailment becomes larger.

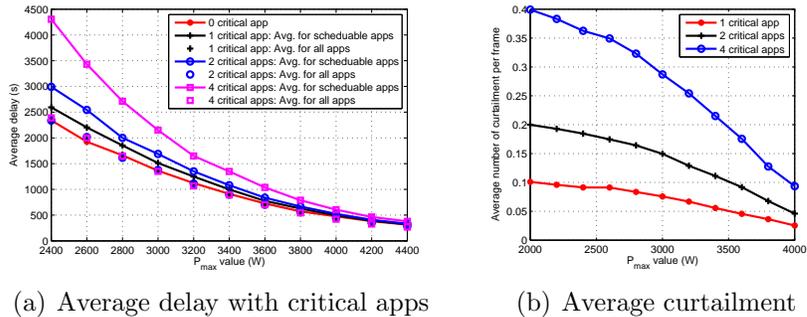


Figure 3.7: Critical appliances effects. ( $N = 15$ ,  $\lambda = 1/900$ , and  $\mu = 1/3000$ )

## (2) Load Scheduling Results

We plot the power consumption of the proposed scheme using an optimized  $\mathcal{P}_{\max,t}$  obtained from (3.23). We employ the price profile of 100 consecutive days as the day-ahead location marginal price (LMP) in the New York City zone from archival data (09/01/2011–12/31/2011) of New York ISO [36]. Given the wind power data, the specified upper and lower bounds ( $\mathcal{U}_t$ ,  $\mathcal{L}_t$ ) and the total power consumption constraint ( $\mathcal{E}_{t_0}$ ), the plot of the optimized  $\mathcal{P}_{\max,t}$  for a day is shown in Figure 3.8(c). For comparison, we also show the plot of  $\mathcal{P}_{\max,t}$  which does not consider the wind

power in Figure 3.8(b), i.e.,  $\tilde{W}_t = 0$  in the optimization problem (3.23). The wind power and electricity price profile for that day are shown in Figure 3.8(a). The following comments are in order: (i) The optimization enables  $\mathcal{P}_{\max,t}$  to be lower during high-price periods and vice versa, which improves the customers' savings in energy cost; (ii) during periods of wind power,  $\mathcal{P}_{\max,t}$  is higher, which encourages customers to use the cheaper wind power, alleviating congestion in appliance access and thereby lowering average delay for each appliance.

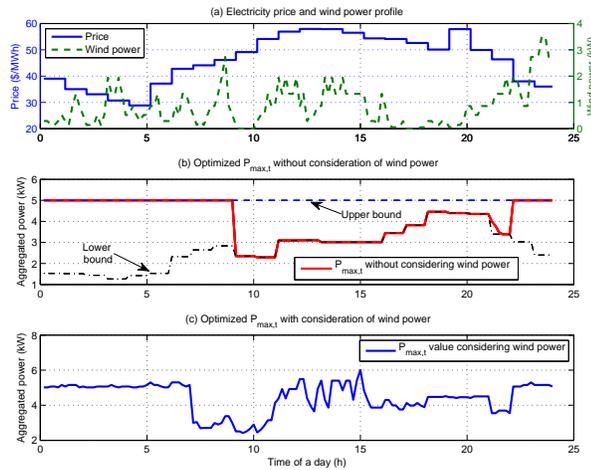


Figure 3.8: Optimized power budget  $\mathcal{P}_{\max,t}$  for a day

We note that the benefits discussed above depend on how closely the actual demand using the proposed scheduling scheme approaches to the target value  $\mathcal{P}_{\max,t}$ . Figure 3.9 shows the scheduling results of our scheme using the power budget obtained in Figure 3.8. Here, we vary the mean arrival rate of appliances based on the time-of-day. Specifically, we assume an arrival rate of  $\lambda = 1/1800$  during the hours [1 : 7];  $\lambda = 1/900$  during the hours [8 : 17] and [23 : 24]; and  $\lambda = 1/200$  during the hours [18 : 22]. We see that, for both the cases of with and without wind power, the scheduled total power consumption is below and close to the target value  $\mathcal{P}_{\max,t}$ . The proposed scheme enables powering-on of appliances to be scheduled as long as the power budget is available. Further, the max-power ACM guarantees the

smallest gap between actual consumption and the target power  $\mathcal{P}_{\max,t}$ .

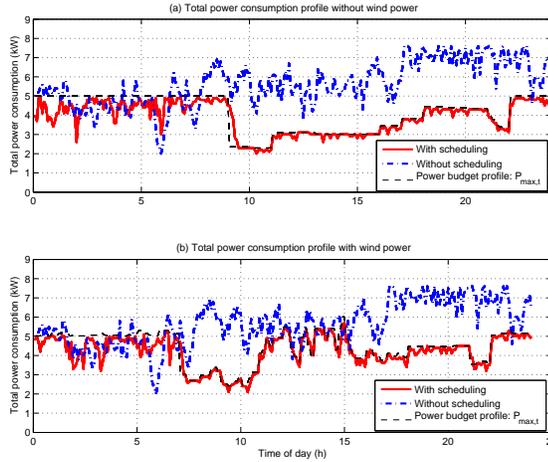


Figure 3.9: Total power consumption under our scheme using  $\mathcal{P}_{\max,t}$  with/without wind power. ( $N = 25$ ,  $\mu = 1/3600$  and random ACM)

Next, in Figure 3.10(a), we compare the cost of buying electricity from the grid over 100 days under the proposed scheme, and the non-scheduling policy; both policies consider the influence of wind power. We see that, with and without wind power the cost saving using the proposed scheme is 34.11% and 33.73%, respectively, better compared to the non-scheduling policy. The reason can be attributed to the fact that optimized  $\mathcal{P}_{\max,t}$  values encourage power usage during low price periods, and the proposed protocol enables appliances to make the best effort to approach the target without exceeding it. We also see that wind power integration can further lower the cost by about 7.39% under our scheme, demonstrating the benefits of using distributed renewable energy at the residential level. The saving is further increased when wind power accounts for a larger portion of the overall supply in-home.

Lastly, Figure 3.10(b) compares the average delay of our scheme with and without consideration of the wind power. We see that the average delay with wind power is slightly lower (6%) than that without wind power. This is because the

optimized  $\mathcal{P}_{\max,t}$  can be larger at the time when there is wind power; larger power budget incurs less congestion and lower average delay.

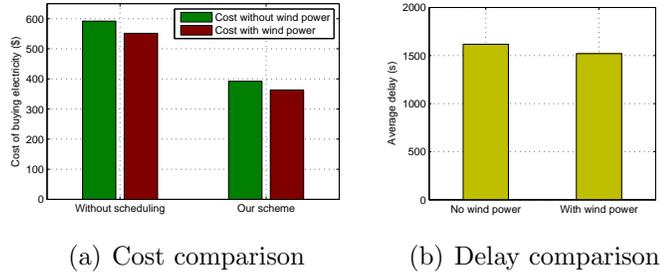


Figure 3.10: Comparison of electricity cost and average delay.

## 3.6 Chapter Summary

In this chapter, we propose a joint access and scheduling scheme for appliances operating over a consumer-premise communication network. The scheme could be incorporated in various DR schemes, such as smart pricing-based or DLC-based. The scheme considers both schedulable appliances, whose powering-on can be delayed, and critical appliances that require real-time power usage. We model the evolution of the protocol as a two-dimensional Markov chain and derive the average delay of an arbitrary appliance. We formulate an optimization problem for the EMC, enabling it to compute the target power level for the home while incorporating effects of price variations and local wind power uncertainty. Simulations verified the analysis and showed that the proposed scheme lowers electricity costs for customers.

# Chapter 4

## Reducing Peak Demand by Real-time Pricing

### 4.1 Introduction

In this chapter, we focus on demand response in the retail market, and study how scheduling capability of certain types of flexible appliances combined with real-time pricing (RTP) interact with objectives of the electric utility. RTP is considered as a very direct and efficient approach for DR [50]. With RTP, the electricity utility announces electricity prices on a rolling basis, i.e., the price for a given time period (e.g., an hour) is determined and announced before the start of the period (e.g., 15 minutes beforehand). With the development of smart metering technologies [3], which will enable secure, reliable, real-time, and two-way information exchange between consumers and their electricity service providers, these RTPs can be provided to consumers multiple times a day, hour, or even second. To handle the resulting data volume and decision making velocity, consumers will rely

on energy management controllers (EMCs) [9], which are devices or programs that use electricity prices and user preferences to modify power usage across a home or building. From the service provider’s perspective, providing high frequency pricing updates to EMCs will enable better load shaping and thus better matching of volatile supply and demand. From the consumer’s perspective, RTP will provide new opportunities to lower rates, provided that they (i.e., EMCs) make smart usage decisions.

Here we propose a smart RTP-based power scheduling scheme for residential power usage using a Stackelberg game model. In this model, the electricity provider plays the leader level game by setting the real-time price and the consumer’s EMC, which schedules appliances in a home, plays the follower level game. The sequential equilibrium is obtained through a two-way information exchange enabled by some underlying communication infrastructure (e.g., the smart metering network). Results show that our scheme can alleviate peak load and reduce the variance between the actual demand and desired demand profile set by the service provider/utilities, which implies substantial cost and stability benefits for the service provider. At the same time, we show that our approach can enable benefits for consumers as well through reduced electricity bills.

The results and analysis in this chapter differ from the related work in several aspects. In [44], consumers (i.e., EMCs) make decisions on their hourly aggregate power consumption using current RTPs. Since individual devices have start and end times that can be deterministic or random, or they may cycle on and off, hourly aggregate consumption values may not be easily mapped to how each device or appliance is operated. By contrast, in our scheme, EMCs schedule power consumption on an appliance-by-appliance basis. In [51, 52], the RTP scheme assumes that the power consumption of each appliance can be modified to optimize

power allocation during each hour. This assumption is suitable for a few devices (e.g., electric vehicle batteries), but most end-use devices do not offer that level of flexibility. On the other hand, several devices do offer flexibility regarding when they are operated, and it is this ‘schedulability’ of appliances that we exploit to demonstrate the power of RTPs and EMCs to collectively match demand with supply. In [53], a bilevel model is proposed to convert the interaction between a large customers with flexible loads and an electricity service provider into a linear programming problem, while in our approach the price is updated when each appliance starts to consume power in a real-time manner.

The remainder of this chapter is organized as follows. Section 4.2 formulates and analyzes the proposed Stackelberg game. Section 4.3 proposes algorithms for our RTP-based power scheduling scheme. Section 4.4 presents numerical results, and we summarize our discussion in Section 4.5.

## 4.2 Stackelberg Game Model Analysis

We consider a residential power system which consists of a service provider and several consumers in a neighborhood, as shown in Figure 4.1. The service provider buys electricity from the wholesale market and sells it to consumers. The EMC in each home interacts with the service provider through an underlying two-way communication network (e.g., the smart metering infrastructure). The EMC coordinates power use among in-home smart appliances; of particular interest in our model, it schedules the time of use of *schedulable* appliances within the home.

Time is divided into slots for scheduling and RTP updates. Let  $\mathcal{T}$  denote the set of time slots in a given time horizon, and  $\mathcal{N}$  denote the set of consumers, where

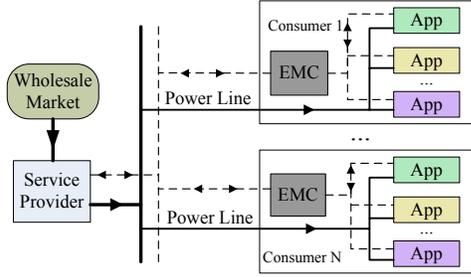


Figure 4.1: The interaction between consumers and electricity service provider

$T \triangleq |\mathcal{T}|$ ,  $N \triangleq |\mathcal{N}|$ . For each consumer  $n \in \mathcal{N}$ , let  $\mathcal{A}_n$  denote the set of schedulable appliances in the home. For appliance  $a \in \mathcal{A}_n$ , the operation duration is denoted by  $l_{n,a}$ , which will be set at the time of request, and the power usage for this duration is  $c_{n,a}$  kW.

We employ a Stackelberg game which is divided into two levels: the service provider plays the *leader level* game and EMCs play the *follower level* game.

#### 4.2.1 EMC/follower level decisions

The EMC aims to minimize the cost to the consumer for an appliance's usage. Its action is to determine the optimal start time  $s$  for a schedulable appliance that was requested to turn on at time slot  $t_0$ . Delaying the appliance to a cheaper price period will save money, but this delay itself also incurs an inconvenience cost. We assume each time slot of delay for appliance  $a \in \mathcal{A}_n$  implies a cost of  $\psi_{n,a}$  dollars. We assume the consumer specifies a maximum allowable delay of  $d_{n,a}$  time slots and the EMC is required to schedule all appliances within the horizon  $\mathcal{T}$ . Given the price vector  $\mathbf{\Pi} = \{\pi_1, \pi_2, \dots, \pi_T\}$  for the time horizon, the optimal scheduled start time  $s^*$  is obtained by solving the following optimization:

$$\begin{aligned}
\min_s \quad & (s - t_0)\psi_{n,a} + \sum_{r=s}^{s+l_{n,a}} \pi_r c_{n,a} \\
\text{s.t.} \quad & t_0 \leq s \leq t_0 + d_n, \quad s + l_{n,a} \leq T.
\end{aligned} \tag{4.1}$$

The minus of the objective function above thus forms the utility function for the EMC follower game. In the objective function of (4.1), the first term represents the delay cost incurred by waiting a time of  $s - t_0$  slots to turn the appliance on. The second term represents the money charged to the consumer by using this appliance. Note that this minimization problem can be solved in  $O(d_{n,a} \cdot l_{n,a})$  time, which is of polynomial complexity.

### 4.2.2 Service provider/leader level decisions

The service provider sets the retail price  $\pi_t$ , which is the sum of the wholesale price  $\phi_t$  and the price gap  $\epsilon_t$ . In our model, the wholesale price affects EMCs' scheduling so that peak load is reduced while the price gap enables the actual load to be more close to the planned supply. The wholesale price  $\phi_t$  is defined as  $\phi_t = C_t(q_t)/q_t$ , where  $q_t$  in kW is the planned supply load for time slot  $t$  and forms the vector  $\mathbf{Q} = [q_1, q_2, \dots, q_T]$  for the time horizon.  $C_t(q_t)$  is the *cost function* [52], which we assume is an increasing and strictly convex function of  $q_t$ ; thus  $\phi_t$  is higher during high load periods than during low load periods. We can choose the cost function as  $C_t(q_t) = \alpha q_t^2$ , where the coefficient  $\alpha$  converts the cost to a monetary value. In this case,  $\phi_t = C_t(q_t)/q_t = \alpha q_t$ .

The price gap  $\epsilon_t$  is designed to influence the difference between the actual demand and the available supply. The service provider maintains a real-time load vector  $\mathbf{Z} = [z_1, z_2, \dots, z_T]$  in kW which tracks the aggregated load for the time

horizon  $\mathcal{T}$ . We design  $\epsilon_t$  such that it is proportional to a function  $g(q_t, z_t, w)$ , which decreases with  $\delta_t = q_t - z_t$ , i.e., the larger  $\delta_t$ , the lower price gap  $\epsilon_t$  so that the EMC is more willing to schedule the appliance to operate during this period, and vice versa. In this paper, we adopt  $g(q_t, z_t, w)$  as:<sup>1</sup>

$$\epsilon_t \propto g(q_t, z_t, w) = \begin{cases} \frac{1}{(q_t - z_t)^w}, & \text{if } q_t > z_t \\ (z_t - q_t)^w, & \text{if } q_t < z_t \\ 1, & \text{if } q_t = z_t; \end{cases} \quad (4.2)$$

where  $w$  is an incentive factor.

We assume a constraint for the price gap which can be seen as the effect of either market competition or price caps set by some regulatory body. For an appliance's request at  $t_0$ , this constraint is  $\sum_{t=t_0}^T \epsilon_t = M_{t_0} = \sum_{t=t_0}^T \epsilon_0$ , where  $\epsilon_0$  is the comparable constant price gap in some (alternate) fixed rate pricing scheme. The price gap  $\epsilon_t$  is expressed in terms of  $g(q_t, z_t, w)$  as:

$$\epsilon_t = \frac{g(q_t, z_t, w)}{\sum_{t=t_0}^T g(q_t, z_t, w)} M_{t_0}. \quad (4.3)$$

The retail price  $\pi_t$  is then:

$$\pi_t = \phi_t + \epsilon_t = \alpha q_t + \frac{g(q_t, z_t, w)}{\sum_{t=t_0}^T g(q_t, z_t, w)} M_{t_0}. \quad (4.4)$$

Given the scheduled start time  $s$ ,  $c_{n,a}$  and  $l_{n,a}$ , the service provider forms a power consumption vector  $\mathbf{P}_{n,a} = [p_{n,a}^1, p_{n,a}^2, \dots, p_{n,a}^t, \dots, p_{n,a}^T]$  in kW for this ap-

---

<sup>1</sup>We round  $q_t$  and  $z_t$  to integer numbers in kW.

pliance, where

$$p_{n,a}^t = \begin{cases} c_{n,a}, & t \in [s, s + l_{n,a}) \\ 0, & t \in \mathcal{T} \setminus [s, s + l_{n,a}). \end{cases} \quad (4.5)$$

The real-time load vector is then updated as

$$\mathbf{Z}' = \mathbf{Z} + \mathbf{P}_{n,a} = [z'_1, z'_2, \dots, z'_T]. \quad (4.6)$$

The utility function of the service provider is defined as the gross profit,  $GP$ , which is  $\sum_{t=1}^T \pi_t \cdot p_{n,a}^t$ , minus the cost of this usage to the provider. This cost has two parts: One cost  $C_e$  comes from purchasing electricity for this appliance usage from the wholesale market, i.e.,  $C_e = \sum_{t=1}^T \phi_t \cdot p_{n,a}^t$ ; the other cost  $C_m$  is due to the “mismatch” between the actual load and planned supply caused by this appliance, i.e.,  $C_m = \beta \left[ \sum_{t=1}^T (q_t - z'_t)^2 - \sum_{t=1}^T (q_t - z_t)^2 \right]$ , where  $\beta$  is a coefficient that converts the cost to a monetary value.

The service provider will maximize its utility by choosing incentive factor<sup>2</sup>  $w$ ; this optimization problem can be written as

$$\max_w \sum_{t=1}^T \epsilon_t p_{n,a}^t - \beta \left[ \sum_{t=1}^T (q_t - z'_t)^2 - \sum_{t=1}^T (q_t - z_t)^2 \right]. \quad (4.7)$$

### 4.2.3 Equilibrium of the Stackelberg game

For each  $w$  chosen by the provider, there is a corresponding retail price vector  $\mathbf{\Pi}^w$ . Using these price vectors, the EMC can determine an optimal start time  $s^w$  for the appliance under consideration. The sequential equilibrium for this game is

---

<sup>2</sup>There is a trade-off in selecting  $w$ : Larger  $w$  provides more incentive for shifting the current appliance’s load to a period with a larger  $\delta_t$  value (thus reduces the mismatch cost  $C_m$ ). But larger  $w$  also reduces the money charged to the consumer, and therefore the gross profit.

the set of values  $(w^*, s^*)$  such that the optimal start time  $s^*$  corresponding to the price vector  $\mathbf{\Pi}^{w^*}$  maximizes the provider's utility function, i.e. implying that  $w^*$  is optimal.

Using backward induction [54], the sequential equilibrium can be solved analytically. On the other hand, we can simplify the optimization problem in (4.7) by discretizing the set of feasible values of  $w$  as  $\mathcal{W} = \{w_1, w_2, \dots, w_W\}$ , where  $W \triangleq |\mathcal{W}|$ . Then the optimization problem in (4.7) can be rewritten as

$$\max_{w \in \mathcal{W}} \sum_{t=1}^T \epsilon_t p_{n,a}^t - \beta \left[ \sum_{t=1}^T (q_t - z'_t)^2 - \sum_{t=1}^T (q_t - z_t)^2 \right]. \quad (4.8)$$

In this simplified model, the sequential equilibrium  $(s^*, \hat{w}^*) \in \mathcal{S} \times \mathcal{W}$  can be achieved through information exchange taking advantage of the two-way communication network.

### 4.3 RTP-based Power Scheduling Scheme

In a neighborhood scenario which contains  $N$  EMCs and thus  $\sum_{n=1}^N |\mathcal{A}_n|$  appliances, the service provider can use its pricing scheme to influence the aggregate load from all these scheduled appliances. For example, in day-ahead pricing (where the prices for each time period in the next day are set a day in advance), each home's EMC can schedule appliance usage to avoid high load (and thus high cost) periods. However, since local neighboring EMCs may make similar usage decisions, the aggregate load for the neighborhood may not experience peak load reduction since a "rebound" peak may appear during what was supposed to be a low load (i.e., low cost) period. Our scheme, presented here as set of interactive algorithms

(Algorithm 1 for the provider, Algorithm 2 for the EMC<sup>3</sup>), alleviates this problem by updating the price according to the real-time load vector  $\mathbf{Z}'$ . In this way, a new appliance's scheduled use will affect future prices and thus the scheduling decisions for future appliances.

---

**Algorithm 1** Executed by the service provider

---

```

1: Initialization.
2: repeat
3:   if receive request signal from EMC  $n$  for app  $a$  then
4:     for  $w=w_1$  to  $w_W$  do
5:       Compute the price vector  $\mathbf{\Pi}^w$  using (4.4).
6:       Send  $\mathbf{\Pi}^w$  to EMC  $n$ .
7:       for all start time  $s^{*,w}$  received do
8:         Solve (4.8) to find the optimal  $\hat{w}^*$ .
9:       end for
10:      Send the  $\hat{w}^*$  to EMC  $n$ , and update  $\mathbf{Z}$  as (4.6).
11:    end for
12:  end if
13: until The end of the day

```

---



---

**Algorithm 2** Executed by EMC  $n$

---

```

1: Initialization.
2: if consumer  $n$  has a request for appliance  $a$  at  $t$  then
3:   Send the request signal to service provider.
4:   for all price vector  $\mathbf{\Pi}^w$  received do
5:     Solve (4.1) to find optimal  $s^{*,w}$  for each  $w$ .
6:     Send  $s^{*,w}$  to the service provider.
7:   end for
8:   if receive the optimal  $\hat{w}^*$  from provider then
9:     Select  $s^* = s^{*,\hat{w}^*}$  and schedule the appliance  $a$ .
10:  end if
11: end if

```

---

<sup>3</sup>Both provider and EMC can compute and receive data simultaneously.

## 4.4 Numerical Results

We simulate a neighborhood consisting of 80 consumers, and each consumer has 3 schedulable appliances, i.e., dishwasher, clothes dryer and clothes washer, controlled by the EMC. We assume  $l_{n,a}$  has an exponential distribution with mean of  $\bar{l}_{n,a}$ . We assume homogeneous end-use consumers, i.e.,  $c_{n,a} = c_a$ ,  $d_{n,a} = d_a$ ,  $\psi_{n,a} = \psi_a$ , and  $\bar{l}_{n,a} = \bar{l}_a$ . In this simulation, we further assume a time horizon of one day, starting from 7:00 AM until the 7:00 AM the next day. Appliances have an on-peak period from 5:00 PM to 8:00 PM, during which time they are requested by consumers with higher probability. We divide time into 10 minute scheduling/pricing slots and use the appliance parameters given in Table 4.1.

Table 4.1: Appliances' parameters for each home

Appliance	$c_a$ (kW)	$d_a$ (hr)	$\psi_a$ (\$/hr)	$\bar{l}_a$ (hr)
Dishwasher	1.8	6.0	0.10	3.0
Clothes Dryer	3.4	4.0	0.25	1.0
Clothes Washer	0.4	2.0	0.40	0.5

### (1) Benefits to the service provider

Figure 4.2 compares non-scheduling aggregate demand for the neighborhood to the load using our RTP-based scheduling scheme. For reference, we also plot the planned supply curve. Without RTP-based power scheduling, the peak load is 102.2 kW at 6:50 PM, while our scheme has a 28.9% lower peak load of 72.8 kW. At the same time, our scheduled consumption curve is much closer to the planned supply curve, i.e., the deviation of the non-scheduled demand curve to the supply curve is 28.3%, which reduces to 16.1% under our scheme.

In Figure 4.2 we also compare the demand curve for our scheme with the demand curve using day-ahead pricing, where each local EMC makes decisions

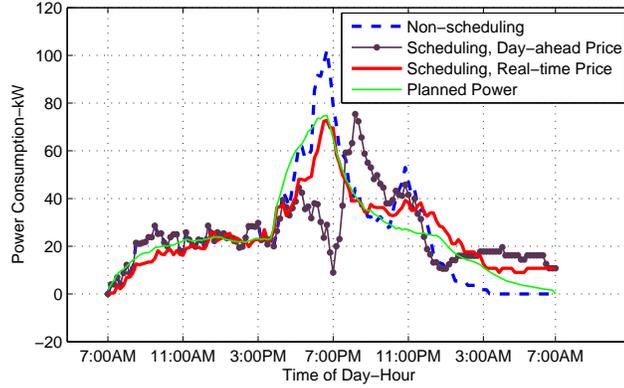


Figure 4.2: One-day power usage comparison

based on the same day-ahead prices. From the plot, we can see that at 6:50 PM, where there is an original peak load, both our real-time pricing and the day-ahead pricing schemes reduce demand by 28.9% and 71.7%, respectively. However, at 8:20 PM, the demand curve for day-ahead pricing exceeds the planned supply load by 81.7%, thereby producing a “rebound” peak. In contrast, our scheme has no such rebound effects.

## (2) Benefits to consumers

Next, we examine the benefits of our proposed scheme for consumers. To do so, we plot in Figure 4.3, the costs (using the objective function in (4.1)) for 10 random consumers for our scheme, the non-scheduling approach and scheduling based on day-ahead prices. Compared to the non-scheduled curve, consumers see a 9.46% reduction in cost for our scheme; while for the day-ahead prices scheme, cost saving is only 5.84%. When considering only electricity costs (i.e., electricity bills), the savings of our proposed scheme is 24.22% over the non-scheduled case, compared to 14.63% savings for the day-ahead scheme.

Our simulation results further show that these cost reductions for the consumers

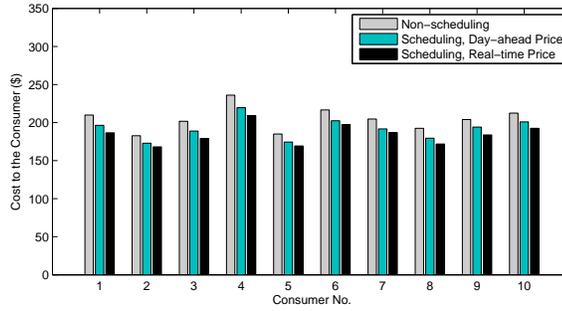


Figure 4.3: Money saving evaluation (100 days)

are consistent across the neighborhood of homes. That is, based on a 100-day average, we observed that the standard deviation in the electricity bills across the 80 homes is within 2.5% of the average electricity bill for the same consumption level. This demonstrates the fairness of our pricing scheme from the consumer’s perspective.

## 4.5 Chapter Summary

In this chapter, we propose a RTP-based power scheduling scheme as a demand response mechanism for residential electric power consumption. A Stackelberg game model is formulated to analyze the interaction between a consumer’s EMC and the service provider. Our scheme can reduce peak load and the mismatch between actual load and planned supply, while avoiding a rebound peak. Using EMCs, consumers can exploit the proposed real-time prices to reduce their electricity bills.

# Chapter 5

## Matching Demand and Supply: A Distributed Direct Load Control Approach

### 5.1 Introduction

In this chapter, instead of the pricing-based method for DR at the retail market level, we focus on an alternative direct load control (DLC) approach, and an innovative distributed direct load control scheme is proposed here to better match demand and supply. The approach utilizes the underlying two-layer communication networks, as shown in Figure 5.1. Specifically, the lower-layer network is within each building, where the energy management controller (EMC) schedules operation of appliances upon request according to a local power consumption target via wireless links. The upper-layer network consists of EMCs in a region of which demand is served by a load aggregator. The load aggregator wants the ac-

tual aggregated demand over this region to match a desired aggregated demand profile determined day-ahead, which is viewed as supply from customers' perspective. Our approach utilizes the average consensus algorithm to distribute portions of the desired aggregated demand to each EMC in a decentralized fashion. The allocated portion corresponds to each building's local power consumption target which its EMC then uses to schedule the in-building appliances. The result will be an aggregated demand over this region that more closely reaches the desired demand.

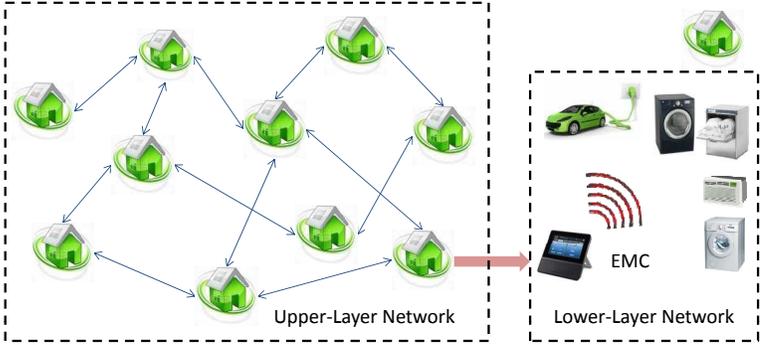


Figure 5.1: Two-layer communication-based control structure

Current DLC programs provided by utilities are usually contract-based, i.e., by signing up for the contract, customers give utilities the option to remotely shut down appliances during high-demand periods or power supply emergency, and receive credit on electricity bills for this participation. Examples of this kind of DLC program include Contracted Direct Load Control by Wisconsin Public Service [7] and Distribution Load Relief Program by conEdison [8]. Several limitations exist for these types of DLC programs. Firstly, they are only for emergency cases so that they do not fully exploit the operational flexibilities of appliances which have greater and longer-term potential for balancing supply and demand. Secondly, when there are a large number of customers, there are huge computational and

communication burdens placed on the centralized controller. Another limit is due to privacy concerns that customers may have since their usage profile is exposed each time an individual appliance is remotely controlled by the central controller.

Whereas there are many existing studies that examine pricing-based approaches for the retail market level DR (e.g., see [42], [44], [55], [56] and references in Chapter 4), there are much fewer papers that look at how direct load control may be conducted under the emerging smart grid infrastructure. Alizadeh et al. [57] proposed a digital direct load control method that categorizes appliances into different queues according to their power consumption request profile, so control of the deferrable load is equivalent to management of the queue parameters, and the control is conducted in a centralized way. [58] proposed a *ColoredPower* distributed algorithm with a probability control method for each appliance such that the switch-on/off decision is based on the probability reflecting the overall supply-demand conditions. However, some parameters of individual appliances still need to be transmitted to the central controller, which incurs a communication burden in larger system.

Our approach has the following advantages compared to the existing methods for large-scale residential DR implementation:

*a. Low complexity for communication and computation:* The distributed algorithm in the upper-layer network that sets local power consumption targets involves only local communications between neighboring EMCs. This approach is therefore of low complexity compared to the centralized method which requires links between each EMC and the centralized controller. Scheduling is conducted in each building by its EMC which incurs lower computation complexity than if the centralized controller made such decisions for each appliance in each building

over the larger region.

*b. Powerful load shaping tools to alleviate the mismatch between demand and supply:* By dividing the desired demand among each building's local power consumption target and designing EMCs that exploit load flexibilities to closely meet these targets, the mismatch between the actual aggregated demand and the desired demand can be alleviated.

*c. Privacy protection for customers:* In our approach, only the aggregated power consumption information of customers is visible in the upper-layer network. Thus, individual appliance usage profile of customers is kept private.

*d. Fairness:* In the proposed scheme, the local power consumption targets for each building are set according to DR resources each customer can provide, with no bias to any individual customer.

*e. Non-intrusiveness for appliance operation:* Our approach integrates non-intrusiveness features for appliance scheduling. This includes options for overriding the control of the EMC, a scheme for preventing frequent ON/OFF switch, and for guaranteeing operation deadlines.

*f. Reliable and robust:* Since local power consumption targets are set in a distributed manner, our approach does not rely on a central controller for the load aggregator, which implies our scheme is more reliable and robust.

The remainder of this chapter is organized as follows. Section 5.2 describes the problem and proposes our distributed DLC architecture. Section 5.3 discusses the distributed demand target allocation in the upper-layer EMC network. Section 5.4 discusses the lower-layer communication and control scheme used by the EMC in each building. Section 5.5 provides numerical results for our approach and we

conclude our discussion in Section 5.6.

## 5.2 Problem Formulation

We consider a region with  $B$  buildings whose power consumptions are served by a load aggregator. The set of buildings is denoted by  $\mathcal{B}$ . Each building  $i \in \mathcal{B}$  has a set of appliances which is denoted by  $\mathcal{K}_i$ . Among these appliances, some are flexible in operation such that they can be switched on either when requested or possibly some time later. The aggregator serving this region has a desired demand  $Z_\tau$  for each time slot  $\tau$ , where the unit of  $\tau$  can be one hour or several minutes depending on the granularity of the design. The desired demand  $Z_\tau$  can be the day-ahead cleared demand bid that the load aggregator submitted into the wholesale electricity market. In real-time, day-of operations, the load aggregator wants the actual aggregated demand of the  $B$  buildings over this region to be lower than  $Z_\tau$  such that the demand commitment in day-ahead ( $Z_\tau$ ) always fulfills the actual demand; meanwhile the actual aggregated demand should be as close to this  $Z_\tau$  as possible in order to minimize the deviation cost.

Instead of a centralized direct load control method, we propose a two-layer communication-based distributed direct load control scheme, shown in Figure 5.1. Each building is equipped with an EMC which locally schedules operation of appliances within the building via a lower-layer (wireless) network. These EMCs are connected by an upper-layer network without a central controller, i.e., each EMC only communicates to its neighboring EMCs.

The main idea underlying the proposed approach is to allocate the desired demand target value  $Z_\tau$  to each EMC via its local power consumption target. This

target is used by the EMC to control appliance operation within the building. Thus, the actual aggregated demand can approach the target  $Z_\tau$ . To reach this goal, two problems need to be solved:

1) How is the local power consumption target within each building computed utilizing the upper-layer network?

2) Within each building, how does the EMC schedule appliances in real-time to reach the local power consumption target, and how are the two-layer processes integrated into a protocol?

In the following sections, we will address how these two problems are tackled in our proposed approach.

## 5.3 Distributed Demand Target Allocation

Let  $\theta_i^\tau$  denote the local demand target set by the EMC within building  $i$  for time slot  $\tau$ . We design a distributed scheme to compute this  $\theta_i^\tau$  so that scheduling appliances locally within each building can result in aggregated demand that approaches the desired demand  $Z_\tau$ .

### 5.3.1 Allocating demand target ( $Z_\tau$ ) to EMCs

Let  $\tilde{q}_i^\tau$  denote the overall power request from delay flexible appliances within building  $i$  for time  $\tau$ . This information can be collected via the lower-layer network which will be discussed in the Section 5.4. These flexible appliances serve as demand response resources that the EMC can exploit to control overall load of the

building. The total demand from already turned-on appliances, i.e., fixed demand, within building  $i$  for time  $\tau$  is denoted by  $\hat{q}_i^\tau$ . We design  $\eta_\tau$  as the ratio of the total demand target gap,  $Z_\tau - \sum_{i \in \mathcal{B}} \hat{q}_i^\tau$ , to the total demand from flexible appliances,  $\sum_{i \in \mathcal{B}} \tilde{q}_i^\tau$ , i.e.,

$$\eta_\tau = \frac{Z_\tau - \sum_{i \in \mathcal{B}} \hat{q}_i^\tau}{\sum_{i \in \mathcal{B}} \tilde{q}_i^\tau}, \quad (5.1)$$

and the EMC for building  $i$  applies this ratio  $\eta_\tau$  to set its local target power consumption target,  $\theta_i^\tau$ , as

$$\theta_i^\tau = \hat{q}_i^\tau + \eta_\tau \cdot \tilde{q}_i^\tau. \quad (5.2)$$

If each EMC schedules appliances to follow this target value  $\theta_i^\tau$ , then we get

$$\sum_{i \in \mathcal{B}} \theta_i^\tau = \sum_{i \in \mathcal{B}} \hat{q}_i^\tau + \eta_\tau \sum_{i \in \mathcal{B}} \tilde{q}_i^\tau = Z_\tau, \quad (5.3)$$

i.e., the aggregated target demand matches the desired demand  $Z_\tau$  for time  $\tau$ . In this approach, the status of aggregated demand reflected by the ratio  $\eta_\tau$  is applied in the local target demand computation for each EMC. Besides guaranteeing the fixed demand  $\hat{q}_i^\tau$ , the additional part of the local target demand  $\theta_i^\tau$  in building  $i$  is proportional to the demand of delay flexible appliances requested in time  $\tau$ . In other words, this demand target allocation is unbiased to each individual customer, only depending on demand response resources.

If there exists a central controller for the aggregator, all EMCs send the fixed demand  $\hat{q}_i^\tau$  and flexible demand  $\tilde{q}_i^\tau$  to the central controller, which then computes the ratio  $\eta_\tau$  according to (5.1), and the central controller then broadcasts this  $\eta_\tau$

back to each EMC. However, the up-link in this scheme will suffer from high communication complexity when the number of EMCs becomes large; meanwhile, this centralized approach is also vulnerable due to its reliance on the central controller, which can be a single point of failure of the system.

An alternative approach is to employ a distributed method, in which the local demand target  $\theta_i^\tau$  is computed at each EMC side via information exchange with its neighboring EMCs using the upper-layer network shown in Figure 5.1. Although  $\tilde{q}_i^\tau$  and  $\hat{q}_i^\tau$  are different for each building, the summation  $\sum_{i \in \mathcal{B}} \hat{q}_i^\tau$  and  $\sum_{i \in \mathcal{B}} \tilde{q}_i^\tau$  are the same for all EMCs, so they can be viewed as a consensus point that all EMCs reach. In this way, the average consensus algorithm [59] and [60] can be applied to compute the  $\eta_\tau$  iteratively.

### 5.3.2 Distributed consensus algorithm to compute $\theta_i^\tau$

The topology of the upper-layer network of EMCs can be represented by an undirected graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  with the set of nodes  $\mathcal{V} = \{1, 2, \dots, B\}$  and edges  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ . The connectivity of the graph  $\mathcal{G}$  is described by the adjacency matrix  $A$  with its entries as

$$A_{i,j} = \begin{cases} 1, & (i, j) \in \mathcal{E}, \\ 0, & \text{otherwise.} \end{cases} \quad (5.4)$$

The set of neighbors of an EMC  $i$  is denoted by  $\mathcal{N}_i$ , where  $\mathcal{N}_i = \{j \in \mathcal{V} : A_{i,j} \neq 0\}$ . By employing the average consensus algorithm [59] and [60] over a connected graph, each EMC  $i$  updates  $\hat{q}_i^\tau$  and  $\tilde{q}_i^\tau$  iteratively as the following steps ( $k$  as the iteration

index)

$$\hat{q}_i^\tau(k+1) = \hat{q}_i^\tau(k) + \varepsilon \cdot \sum_{j \in \mathcal{N}_i} [\hat{q}_j^\tau(k) - \hat{q}_i^\tau(k)], \quad (5.5)$$

$$\tilde{q}_i^\tau(k+1) = \tilde{q}_i^\tau(k) + \varepsilon \cdot \sum_{j \in \mathcal{N}_i} [\tilde{q}_j^\tau(k) - \tilde{q}_i^\tau(k)]. \quad (5.6)$$

The values of all  $\hat{q}_i^\tau$  and  $\tilde{q}_i^\tau$  will converge to the average value  $\hat{q}^{\tau*}$  and  $\tilde{q}^{\tau*}$ , respectively, with a proper step size  $\varepsilon$ . After the updates converge, the global ratio  $\eta_\tau$  can be computed at each EMC side as:

$$\eta_\tau = \frac{Z_\tau - B \cdot \hat{q}^{\tau*}}{B \cdot \tilde{q}_i^{\tau*}}, \quad (5.7)$$

where the overall desired target demand,  $Z_\tau$ , and total number of EMCs  $B$ , are known to all EMCs beforehand.

The convergence of the average consensus algorithm in (5.5) and (5.6) is related to the graph Laplacian matrix of the network graph  $\mathcal{G}$ , which is denoted by  $L$  such that  $L = D - A$ , where  $D = \text{diag}\{d_1, \dots, d_B\}$  is the degree matrix of  $\mathcal{G}$ , with  $d_i$  being the number of neighbors of node  $i$  with which it can communicate reliably, i.e.,  $d_i = |\mathcal{N}_i|$ . Given this matrix  $L$ , it can be shown that  $L\mathbf{1} = \mathbf{0}$  and  $\mathbf{1}^\text{T}L = \mathbf{0}^\text{T}$ , where  $\mathbf{1} = [1, \dots, 1]^\text{T}$ , and  $\mathbf{0} = [0, \dots, 0]^\text{T}$  with proper size. Additionally,  $L$  is a symmetric positive semi-definite matrix, and for a connected graph, the rank of  $L$  is  $B - 1$  and its eigenvalues can be arranged in increasing order as  $0 = \lambda_1(L) < \lambda_2(L) \leq \dots \leq \lambda_B(L)$  [61]. From [60] and [62], it can be shown that for a time-invariant, connected, undirected network, when  $\varepsilon \in (0, 2/\lambda_B(L))$ , average consensus can be asymptotically achieved. Specifically, the minimal convergence time is obtained when  $\varepsilon = 2/[\lambda_2(L) + \lambda_B(L)]$  [62], i.e., the second smallest eigenvalue of graph Laplacian quantifies the speed of convergence of consensus al-

gorithms.

In regard to convergence speed, [63] shows that a discrete-time consensus is globally exponentially reached with a speed that is faster or equal to  $\kappa_2 = 1 - \varepsilon\lambda_2(L)$  for a connected undirected network. This property helps in designing the network topology to achieve a better trade-off between the infrastructure cost and convergence rate. Figure 5.2 shows four typical network topologies of 20 nodes, where the second largest eigenvalue of graph Laplacian,  $\lambda_2(L)$ , and the convergence speed parameter  $\kappa_2$  of each network is computed and shown in the figure. We

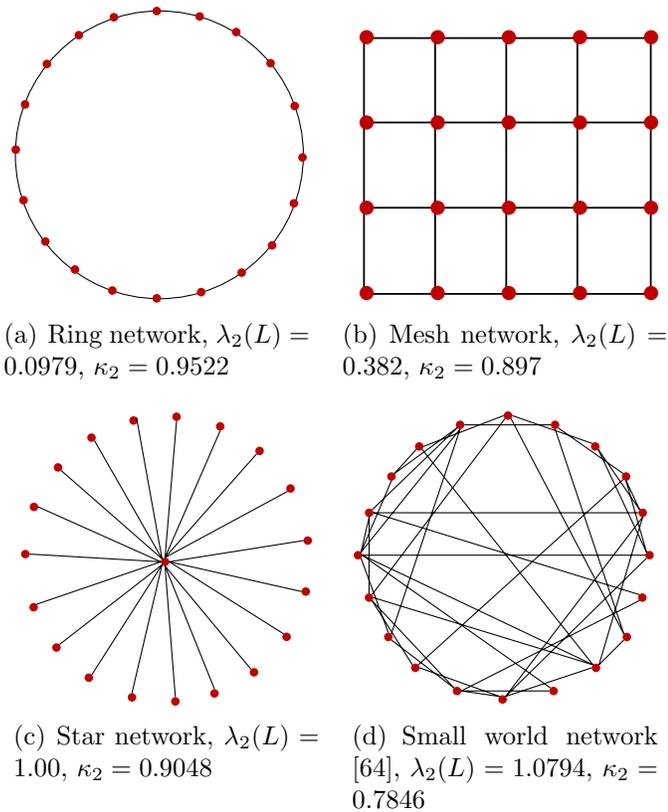


Figure 5.2: Four sample networks with 20 nodes

can see among these four networks, the small-world network has the lowest  $\kappa_2$  values, indicating the fastest convergence speed. This can be verified in Figure 5.3 where the convergence for these four sample networks are illustrated. The

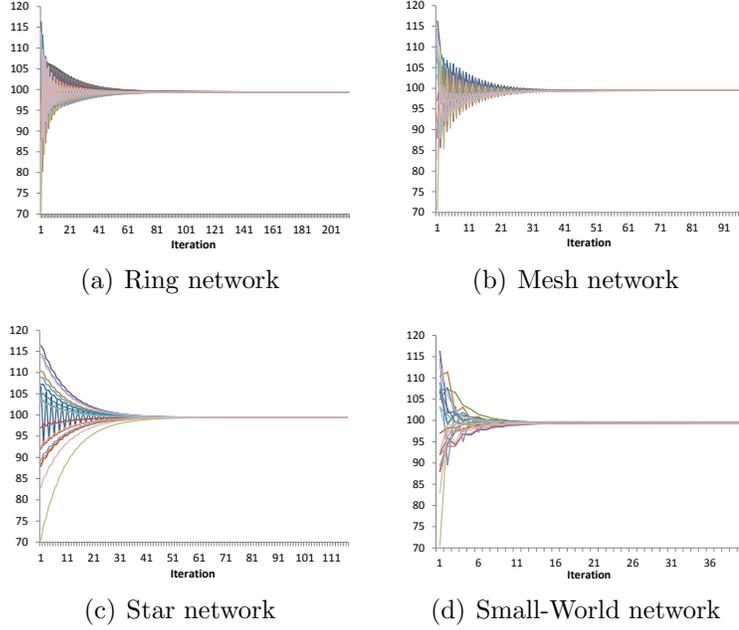


Figure 5.3: Convergence properties for four networks

convergence speed can also help design the parameters of the communication and control protocol to guarantee that the distributed method gives an accurate result; this will be discussed in Section 5.4.

Given the ratio  $\eta_\tau$  from the average consensus algorithm, the local power consumption target  $\theta_i^\tau$  for building  $i$  can be computed locally according to (5.2). Depending on the values of  $\eta_\tau$ , three scenarios need to be considered:

If  $\eta_\tau \leq 0$ , i.e., the desired demand  $Z_\tau$  is less than the aggregated fixed demand, the local power consumption target could be set as the its fixed value,  $\hat{q}_i^\tau$ , by letting  $\eta_\tau = 0$ . In this case, all flexible appliances will not be turned on during that time slot, and a deviation cost will be incurred in order to fulfill the fixed demand.

If  $0 < \eta_\tau \leq 1$ , i.e., the current demand from delay flexible appliances' request is larger than the desired demand less the overall fixed demand, then  $\theta_i^\tau$  can be computed directly using (5.2).

If  $\eta_\tau > 1$ , then the local power consumption target can be set as  $\theta_i^\tau = \hat{q}_i^\tau + \tilde{q}_i^\tau$ , i.e., by letting  $\eta_\tau = 1$ . This represents the best effort that the EMC can do to schedule appliances at time  $\tau$  to reach the target. On the other hand, if battery storage is locally available and can be utilized, this gap can be filled. But this scheme needs to quantify the fairness of the batteries' charging allocation among customers as well as other constraints for batteries; these issues are left for future work.

The distributed algorithm proposed here only requires local communication between neighboring EMCs, and does not rely on a central controller. In addition, since only aggregated information of each building,  $\hat{q}_i^\tau$  and  $\tilde{q}_i^\tau$ , are transmitted (instead of individual appliance information), customer privacy is better protected.

## 5.4 Lower-Layer Communication and Admission Control Scheme

Each EMC  $i$  in the lower-layer network has two tasks: 1) Collect power request information from appliances within the building and update the  $\hat{q}_i^\tau$  and  $\tilde{q}_i^\tau$  for upper-layer consensus algorithm to compute  $\theta_i^\tau$ . 2) Conduct admission control for delay flexible appliances to decide whether they can be turned on to achieve the local power consumption target  $\theta_i^\tau$ . Before providing details on how these two tasks may be executed, we first describe the structure of our two-layer communication-based control approach, as shown in Figure 5.4.

Time is segmented into frames which last as long the slot  $\tau$  used in Section 5.3. Each frame  $\tau$  has three phases:

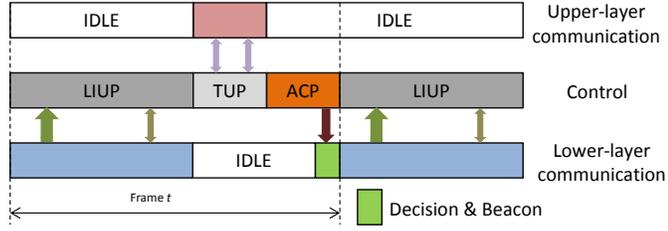


Figure 5.4: Two-layer communication and control protocol

**(1) Load Information Update Phase (LIUP):** During the LIUP, the lower-layer communication interface module receives the power-on requests and power-off requests from appliances in the building. With this data collected, the control module updates the flexible load  $\tilde{q}_i^\tau$  and fixed load  $\hat{q}_i^t$  accordingly. The details of the scheme during this phase is described in Section 5.4.1.

**(2) Target Update Phase (TUP):** During the TUP, the upper-layer communication interface module at each EMC communicates with its neighboring EMCs to execute the average consensus algorithm discussed in Section 5.3 and compute the local demand target value  $\theta_i^\tau$  for frame  $\tau$ . The length of this phase should guarantee the convergence of the average consensus algorithm discussed in Section 5.3.

**(3) Admission Control Phase (ACP):** During the ACP, the EMC's control module makes the admission decision, i.e., decide which delay flexible appliances can be turned on, given the local demand target value  $\theta_i^\tau$ . At the end of the ACP, the lower-layer communication interface module sends a beacon signal to broadcast the admission decision and synchronizes to the next frame. Details of the admission control scheme are described in Section 5.4.2.

Note that in Figure 5.4 the length of the blocks do not show the relative length

of each phase. The duration of each phase is a design parameter of the protocol which can be adjusted according to the implementation scenario.

### 5.4.1 Load information update

To track fixed demand  $\hat{q}_i^\tau$  and flexible demand requests  $\tilde{q}_i^\tau$ , each EMC  $i$  maintains the following sets of appliances in building  $i \in \mathcal{B}$ :

- *Active Appliance Set* ( $\mathcal{A}_i^\tau$ ) contains already powered on appliances during the frame  $\tau$  in building  $i$ .

- *Request Appliance Set* ( $\mathcal{R}_i^\tau$ ) tracks appliances that are requested to power on by the customer and their power-on request messages (PRQMs) are successfully received by the EMC in building  $i$  during frame  $\tau$ . If appliance  $j \in \mathcal{K}_i \setminus (\mathcal{A}_i^\tau \cup \mathcal{R}_i^\tau)$  is requested to power on, it sends a PRQM; upon receiving this message, the EMC updates  $\mathcal{R}_i^\tau$  as  $\mathcal{R}_i^\tau = \mathcal{R}_i^\tau \cup \{j\}$ .

- *Release Appliance Set* ( $\mathcal{X}_i^\tau$ ) tracks appliances that wish to power off and their power release messages (PRLMs) have been successfully received by the EMC in building  $i$  during frame  $\tau$ . Upon receiving the PRLM from an appliance  $j \in \mathcal{A}_i^\tau$ , the EMC updates  $\mathcal{X}_i^\tau$  to  $\mathcal{X}_i^\tau = \mathcal{X}_i^\tau \cup \{j\}$ .

Note that for a delay flexible appliance,<sup>1</sup> whose PRQM (PRLM) is received by the EMC, the appliance is not added to (removed from) *Active Appliance Set* ( $\mathcal{A}_i^\tau$ ) immediately; instead it is added to  $\mathcal{R}_i^\tau$  ( $\mathcal{X}_i^\tau$ ) and waits until the ACP phase to be scheduled on/off by the EMC. Thus,  $\mathcal{R}_i^\tau$  and  $\mathcal{X}_i^\tau$  only contains delay flexible

---

<sup>1</sup>Note that some appliances have delay flexibilities in both turning-on and turning-off, e.g., electric vehicle battery and water heater. Some appliances, however, only provide delay flexibilities regarding when they turn-on, e.g., clothes washer; in such cases, we can treat their turn-off requests via the override option, discussed in Section 5.4.3

appliances. When non-flexible appliances wish to turn on or off, they will be treated as appliances with override option and will be added to (removed from)  $\mathcal{A}_i^r$  immediately; this option will be discussed in Section 5.4.3.

The EMC updates these appliance sets according to the packets (PRQMs or PRLMs) transmitted by appliances. Table 5.1 shows the essential fields in these packets required for our approach. The RRI indicates whether this packet is a PRQM or PRLM packet, and PCV is the power consumption value of the appliance. Other fields (ORI, MON, MSUS, DT) are related to the non-intrusive operation of appliances which will be discussed in Section 5.4.3.

Table 5.1: Fields description in appliance TX packet

ID	Appliance/EMC ID
RRI	PRQM/PRLM indicator (0/1 binary)
ORI	Override indicator (0/1 binary)
PCV	Power consumption value in W
MON	Minimum on time
MSUS	Minimum suspension time
DT	Delay tolerance

The EMC transmits an ACK packet to acknowledge successfully reception of the PRQM/PRLM sent by an appliance. ACK is integrated in the general EMC transmission packet. Table 5.2 shows the fields needed in the EMC transmission packet. For the EMC's ACK transmission, the ACKID field is assigned the appliance's ID and thus acknowledges the success of packet reception from that appliance. Other fields (BCI, OPD, RXD) are valid in the ACP phase which will be discussed in Section 5.4.2.

Carrier sense multiple access with collision avoidance (CSMA/CA), which is widely used in wireless local area networks like IEEE 802.11 [65], is applied for packet transmissions in our scheme. At the end of the LIUP, if an appliance's

Table 5.2: Fields description in EMC TX packet

ID	Appliance/EMC ID
BCI	Beacon packet indicator
ACKID	ACK destination appliance ID, ACKID=0 for broadcast
OPD	Appliances operation (on/off) decision vector, valid for the beacon and broadcast (ACKID=0) packet
RXD	Specify a RX duration in which appliances are allowed to transmit request/release messages, valid for beacon packet

request is not received by the EMC due to collision, it will wait until the next frame to send its request; however, if the duration of LIUP is long and the number of requests is not large, the impact of collisions can be neglected.

At the end of the LIUP phase, with updated set  $\mathcal{R}_i^\tau$  and  $\mathcal{X}_i^\tau$  as well as  $\mathcal{A}_i^\tau$  for frame  $\tau$ , the fixed demand of building  $i$  can be computed as

$$\hat{q}_i^\tau = \sum_{j \in \mathcal{A}_i^\tau \setminus \mathcal{X}_i^\tau} P_{i,j}^\tau, \quad (5.8)$$

where  $P_{i,j}^\tau$  is the power consumption value for appliance  $j$  in building  $i$  during frame  $\tau$ .<sup>2</sup> The requested demand from flexible appliances  $\tilde{q}_i^\tau$  can be computed as

$$\tilde{q}_i^\tau = \sum_{j \in \mathcal{R}_i^\tau} P_{i,j}^\tau. \quad (5.9)$$

## 5.4.2 Admission control mechanism

Given the flexible appliances request set  $\mathcal{R}_i^\tau$  and the local demand target value  $\theta_i^\tau$ , the EMC makes admission decisions in the ACP phase, i.e., chooses which

<sup>2</sup>Note that in this discussion we assume  $P_{i,j}^\tau$  is a fixed constant over the frame  $\tau$ . In reality, the power consumption may vary with time. To this end, the length of the frame needs be carefully designed so that our assumption gives accurate results. For those flexible appliances that have fast variation of power consumption with time, we can also treat the change in power consumption as an additional power-on/power-off request that is handled via the override option. The override option will be discussed in Section 5.4.3.

appliances in  $\mathcal{R}_i^\tau$  to turn on. The aim of the controller is to make the aggregated load in building  $i$  approach  $\theta_i^\tau$  as close as possible, without being over  $\theta_i^\tau$ . We employ a binary decision variable  $x_j$  associated to each appliance  $j \in \mathcal{R}_i^\tau$  indicating whether it can be turned on. The EMC's optimization can be formulated as:

$$\max_{x_j} \sum_{j \in \mathcal{R}_i^\tau} x_j \cdot P_{i,j} \quad (5.10)$$

$$\text{s.t. } \hat{q}_i^\tau + \sum_{j \in \mathcal{R}_i^\tau} x_j \cdot P_{i,j} \leq \theta_i^\tau \quad (5.11)$$

$$x_j \in \{0, 1\}, j \in \mathcal{R}_i^\tau, \quad (5.12)$$

where  $x_j = 1$  indicates that appliance  $j$  is scheduled to turn on. (5.10) – (5.12) is a classical 0 – 1 knapsack problem, which can be solved efficiently using various existing approaches [45].

At the end of the ACP phase, the EMC broadcasts a packet containing the admission results, i.e., it sets the OPD field in Table 5.2 accordingly. This packet also serves as a beacon signal (set BCI=1 in Table 5.2) to specify the transmission/receiving duration (length of LIUP) for the next frame, by setting the RXD field in Table 5.2, during which the power on/off request packets can be sent.

With the decision result  $\{x_j\}$ , *Admitted Appliance Set*  $\mathcal{D}_i^\tau$  can be updated as

$$\mathcal{D}_i^\tau = \{j | j \in \mathcal{R}_i^\tau, x_j = 1 \text{ in } (5.10)-(5.11)\}.$$

The *Active Appliance Set* for the next frame  $\tau + 1$  is updated as

$$\mathcal{A}_i^{\tau+1} = \mathcal{A}_i^\tau \cup \mathcal{D}_i^\tau \setminus \mathcal{X}_i^\tau. \quad (5.13)$$

Note that the solution of (5.10) may result in a gap between  $\hat{q}_i^\tau + \sum_{j \in \mathcal{R}_i^\tau} x_j \cdot P_{i,j}$  and  $\theta_i^\tau$ , in which any appliance in  $\mathcal{R}_i^\tau$  can not be accommodated. This gap will most likely exist at each building; the cumulative effect will produce a gap between the actual scheduled aggregate demand and the desired aggregate demand. To alleviate this gap, we employ a relaxed target  $\tilde{\theta}_i^\tau = \theta_i^\tau + P_i^{\max}$ , where  $P_i^{\max}$  is the maximum value of the appliances' power consumptions within building  $i$  during frame  $\tau$ . With a probability of  $\zeta$ , we assume the EMC uses this  $\tilde{\theta}_i^\tau$  in the admission control optimization. The parameter  $\zeta$  is tuned according to the actual implementation. This practical adjustment allows some buildings to use extra power beyond the local demand target, but all buildings have this opportunity without bias due to the probability  $\zeta$ . Note that this scheme cannot remove the gap totally. Although large values of  $\zeta$  may decrease the gap, there may be a possibility that the aggregated demand will exceed the desired demand; thus the trade-off in values of  $\zeta$  should be carefully considered in actual implementation.

### 5.4.3 Non-intrusive operation for appliances

Our approach exploits operational flexibility of appliances in terms of delaying turning-on. In other words, the control scheme may change the original usage profile of appliances. It is feasible for some kinds of non time-critical appliances such as washing machine and clothes dryer, but appliances with critical time requirement, such as TV, desktop, and emergency devices, are required to turn on immediately once the customer wants to use them. In some cases, even for flexible appliances, customers may not prefer to delay usage for a variety of reasons. The underlying DR scheme should provide such override options for customers even though overriding may affect the incentive rebates that customers receive. Besides

the override option, non-intrusive operation may be needed, i.e., guaranteeing the appliance's task is completed before a deadline set by the customer. Furthermore, this should be done while avoiding frequent appliance ON/OFF switches in the scheduling. In the following, we will describe how non-intrusive operations are integrated into our control scheme discussed above.

### (1) Customer override option

Our control scheme provides customers with override option, e.g., to specify time critical appliances or change request preferences. In such cases, the appliance should be turned on/off as soon as the customer requests; meanwhile the appliance sends a PRQM/PRLM packet with the ORI field set, as shown in Table 5.1, to indicate the request is due to an override option.

If the PRQM packet with an override option from appliance  $j$  is received by the EMC, the *Active Appliance Set* is updated as  $\mathcal{A}_i^\tau = \mathcal{A}_i^\tau \cup \{j\}$ . However, since during this frame  $\tau$ , the demand of the building needs to be kept below the demand target value computed at the end of the last frame, i.e.,  $\theta_i^{\tau-1}$ , this newly added appliance may cause the actual demand to go above this target. In this scenario, a curtailment is needed to ensure the demand target  $\theta_i^{\tau-1}$  bound not being violated.

Define the *Curtailable Appliance Set*  $\mathcal{C}_i^\tau$ , where  $\mathcal{C}_i^\tau \subseteq \mathcal{A}_i^\tau$ , as the set of appliances that can be curtailed in frame  $\tau$ . A curtailment optimization is formulated by the

EMC as:

$$\max_{y_j} \sum_{j \in \mathcal{C}_i^\tau} (1 - y_j) \cdot P_{i,j} \quad (5.14)$$

$$\text{s.t.} \quad \sum_{k \in \mathcal{A}_i^\tau \setminus \mathcal{C}_i^\tau} P_{i,k} + \sum_{j \in \mathcal{C}_i^\tau} (1 - y_j) \cdot P_{i,j} \leq \theta_i^{\tau-1} \quad (5.15)$$

$$y_j \in \{0, 1\}, j \in \mathcal{C}_i^\tau, \quad (5.16)$$

where the binary variable  $y_j = 1$  indicates appliance  $j$  in  $\mathcal{C}_i^\tau$  is scheduled to be curtailed. The curtailed appliances will be added to  $\mathcal{R}_i^\tau$  for future scheduling, i.e.,  $\mathcal{R}_i^\tau = \mathcal{R}_i^\tau \cup \{j | j \in \mathcal{C}_j, y_j = 1 \text{ in (5.14)-(5.16)}\}$ .

If the PRLM packet with an override option from appliance  $j$  is received by the EMC, the *Active Appliance Set* is updated as  $\mathcal{A}_i^\tau = \mathcal{A}_i^\tau \setminus \{j\}$ . In this scenario, the gap between the actual demand  $\sum_{j \in \mathcal{A}_i^\tau} P_{i,j}$  and the target value  $\theta_i^{\tau-1}$  will increase, so that some additional appliances in  $\mathcal{R}_i^\tau$  may be accommodated and turned on.

Define the *Admissible Appliance Set*  $\mathcal{S}_i^\tau$ ,  $\mathcal{S}_i^\tau \subseteq \mathcal{R}_i^\tau$  as the set of appliances that can be admitted to power on at this moment. An admission optimization is formulated for the EMC as:

$$\max_{z_j} \sum_{j \in \mathcal{S}_i^\tau} z_j \cdot P_{i,j} \quad (5.17)$$

$$\text{s.t.} \quad \sum_{k \in \mathcal{A}_i^\tau} P_{i,k} + \sum_{j \in \mathcal{S}_i^\tau} z_j \cdot P_{i,j} \leq \theta_i^{\tau-1} \quad (5.18)$$

$$z_j \in \{0, 1\}, j \in \mathcal{S}_i^\tau, \quad (5.19)$$

where the binary variable  $z_j = 1$  indicates appliance  $j$  in  $\mathcal{S}_i^\tau$  is scheduled to turn on and be added to  $\mathcal{A}_i^\tau$ , i.e.,  $\mathcal{A}_i^\tau = \mathcal{A}_i^\tau \cup \{j | j \in \mathcal{S}_j, z_j = 1 \text{ in (5.17)-(5.19)}\}$ .

For both cases, after the decision is made, the EMC (via lower-layer communication interface module) broadcasts an ACK packet in which the BCI and ACKID fields, as shown in Table 5.2, are set to 0 indicating non-beacon and broadcast packet, and the OPD field is set to the admission decision vector to inform all appliances. Upon receiving this broadcast ACK, the corresponding appliances will suspend power consumption (curtailment) or turn on.

Note that the override action may happen during TUP period, so the local target demand  $\theta_i^T$  computation by the average consensus algorithm may not be accurate in terms of actual demand status. However, this problem is trivial since the duration of TUP can be designed to be very short, e.g., several seconds. In such a case, the override operation of many appliances happen during this phase with negligible probability; the override option during this short time can also be disabled, which is feasible for customers.

## **(2) Preventing frequent ON/OFF switching**

With the override option, some flexible appliances may be shut down temporarily (suspended) or resumed. However, too frequent ON/OFF switching is harmful to appliances. To accommodate this non-intrusive constraint, min-ON and min-OFF parameters are specified for each appliance to indicate the minimum power-on time after it is turned on and minimum off time after it is suspended. These two parameters are passed to the EMC by the MON and MSUS fields, respectively, of the PRQM packet shown in Table 5.1.

Accordingly, there are ON and OFF timers for each appliance in the EMC's control module to record the ON/OFF times. When the appliance is turned on, either for the first time or when it resumes from a suspended state, the ON timer

counts down with initial value min-ON. The appliance cannot be added to the *Curtailable Appliance Set*,  $\mathcal{C}_i^T$ , until the timer reaches 0; this prevents the appliance from being curtailed and suspended again. Similarly, when the appliance is curtailed, the OFF timer counts down with initial value of min-OFF. The appliance will not be added to the *Admissible Appliance Set*  $\mathcal{S}_i^T$ , until the timer reaches 0. This mechanism helps avoid frequent ON/OFF switching for appliances due to curtailment and admission from the override operation.

### (3) Operation deadline constraint

Although some flexible appliances allow delayed power-on or suspension, an operation deadline is usually needed to guarantee that the appliance's task operation is completed by the customer's desired deadline. Our scheme integrates this requirement by specifying a DT field in the PRQM packet (shown in Table 5.1), which represents the delay tolerance. This delay tolerance can be directly specified by the customer, or if the customer indicates the deadline for the operation completion, the DT can be computed as  $tc_{i,k} - tr_{i,k} - td_{i,k}$ , where  $tc_{i,k}$ ,  $tr_{i,k}$ , and  $td_{i,k}$  are the deadline point, the request time point, and the duration of operation, respectively.

After the EMC receives the PRQM packet containing the DT, a deadline timer is set to the value of DT. When the appliance waits to get access or suspends, the timer counts down. If the appliance is in power-on status, the timer pauses. If the value of the timer is less than min-OFF parameter (MSUS), it can not be scheduled to be suspended. If the timer reaches 0 before the task completion, the appliance powers on as an override operation and continues to consume power without interruption until the task is completed.

## 5.5 Numerical Results

In the simulation, we consider a region with  $B = 400$  buildings over which the corresponding EMCs form a  $20 \times 20$  mesh network. Each building in turn has 40 appliances to schedule, i.e.,  $|\mathcal{K}_i| = 40$ , 20 of which are delay flexible appliances. Since typical residential appliances' power consumption ranges from 50W to 2000W [66], we assume power consumption values in our simulation to be uniformly distributed, i.e.,  $P_{i,k} \sim U(50\text{W}, 2000\text{W})$ ,  $i \in \mathcal{B}$  and  $k \in \mathcal{K}_i$ . We assume the power duration of appliances follows an exponential distribution with mean of  $\mu_{i,k}^\tau$  (i.e., with rate of  $\frac{1}{\mu_{i,k}^\tau}$ ). The request interval (that is, the duration from the time at which appliance becomes idle to the time when the appliance is requested to power on by the customer) is assumed to be exponentially distributed with mean of  $\alpha_{i,k}^\tau$ .

To make aggregate demand in the simulation be more practical, we employ the actual real-time demand data of New York City from NYISO [67] to tune the appliance power-on request interval parameter  $\alpha_{i,k}^\tau$ . Specifically, we denote the base value of this parameter by  $\alpha_{i,k}^0$ , and the tuned value is computed as  $\alpha_{i,k}^\tau = \gamma^\tau \cdot \alpha_{i,k}^0$ , where the ratio  $\gamma^\tau$  is computed in terms of actual real-time demand data  $Q^\tau$  as

$$\gamma^\tau = \left( \frac{1/L \cdot \sum_{\tau=0}^{L-1} Q^\tau}{Q^\tau} \right)^\omega, \quad \omega \geq 1, \quad (5.20)$$

where  $\omega$  is the tuning parameter, and we choose  $\omega = 1$  in our simulation. From (5.20) we can see the higher actual demand results in smaller power-on request interval parameter  $\alpha_{i,k}^\tau$ ; thus the higher demand in the simulation due to more frequent power-on requests. Figure 5.5 shows the aggregated demand of our test system using the tuned parameters (Figure 5.5-b), and the actual average demand

of N.Y.C used for tuning parameter  $\alpha_{i,k}^T$  (Figure 5.5-a) is also shown for comparison. We can see that the simulated demand follows the trend of the actual demand data.

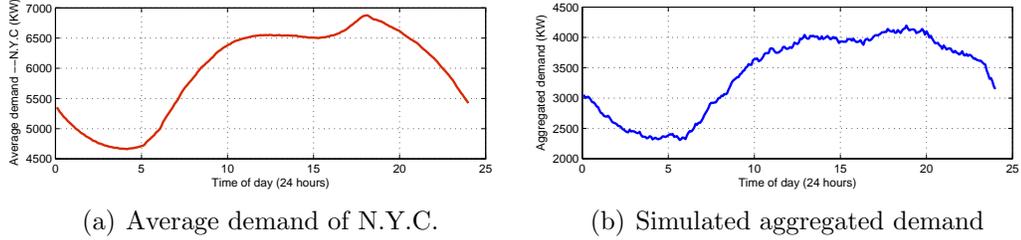


Figure 5.5: Aggregated demand profile for simulation

The step size of the average consensus algorithm in (5.5)–(5.6) can be computed as  $\varepsilon = 2/[\lambda_2(L) + \lambda_B(L)] = 0.2508$ , where the second smallest eigenvalue of graph Laplacian matrix  $\lambda_2(L) = 0.0246$ , which gives the convergence speed parameter  $\kappa_2 = 1 - \varepsilon\lambda_2(L) = 0.9938$ . Assuming the initial standard deviation of  $\{\hat{q}_i^T\}$  and  $\{\tilde{q}_i^T\}$ ,  $i \in \mathcal{B}$  to be  $\varphi_0$ , and the converged deviation to be  $\varphi_c$ , the iteration time  $I$  can be computed as

$$I = \frac{\ln(\varphi_c/\varphi_0)}{\ln(\kappa_2)}. \quad (5.21)$$

Consider the worst case in which  $\{\hat{q}_i^T\}$  and  $\{\tilde{q}_i^T\}$  are polarized for all EMCs, i.e., half of the EMCs have the largest value of  $\sum_{j=1}^{40} P_{i,j}^{\max} = |\mathcal{K}_i| \cdot P_i^{\max}$  and half of the EMCs have zero values, so  $\varphi_0 = |\mathcal{K}_i| \cdot P_i^{\max}/2$ . We set the converged deviation as  $\varphi_c = 0.1$ , so from (5.21) the upper bound of convergence iteration is computed as  $I = 2083$ . Assuming the time of each iteration in the upper-layer network (transmission and processing time) as 0.1ms, we can set the duration of TUP as 0.3s to guarantee convergence.

Figure 5.6 shows the scheduling results for one day under our proposed scheme.

We can see the scheduled aggregated demand (blue curve) is closer to the desired demand (black curve) than the original demand (red curve). With the proposed scheme, the average deviation between the desired demand and actual demand levels over the day decreases by 35.6% compared to that of the original demand.

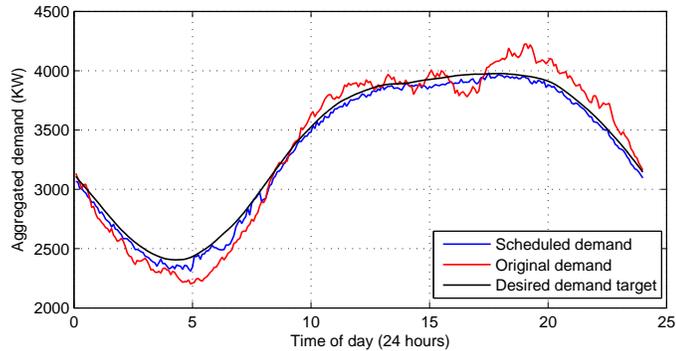


Figure 5.6: Scheduling demand using the proposed approach

## 5.6 Chapter Summary

In this chapter, an innovative two-layer communication-based distributed direct load control approach is proposed for future large-scale residential demand response. The idea is to allocate the overall control task into each building by EMC in a distributed way. We employ average consensus algorithms in the upper layer network, which consists of connections between neighboring EMCs, to allocate target power consumption levels for each building. The EMC in each building then schedules appliances operation according to the local power consumption target. The protocol integrates these two-layer seamlessly and also enables non-intrusive operation of appliances. Numerical results show notable improvement in the system’s ability to match day-ahead demand predictions (i.e., supply) with day-of, actual demand levels.

# Chapter 6

## Market Effects of Deferrable Loads as Demand Response

### 6.1 Introduction

In this chapter, we focus on how demand response may influence the wholesale electricity market. Favored by the FERC Order 745 [15] described in Section 1 stating compensation to DR resources in the wholesale electricity market, we are interested in the market effects of demand side participation of DR aggregators. Specifically, we employ the model proposed in [57], which quantizes load requests from flexible load devices and bundles them into a set of discrete load categories, each category being a queue of energy requests that has a specific load profile. By dispatching appliances activating at a certain time from each queue, the DR aggregator manages a load reservoir formed by the queues of requests. The ability to queue and dispatch loads can be viewed as a special type of generation, from the perspective of balancing supply and demand, with its own generation and

cost. The “generation capacity” for the DR aggregator is largely dependent on the schedulable resource, i.e., the number of flexible loads it controls, their arrival rates, etc. On the other hand, the “generation cost” is the delay inconvenience cost to customers due to queued loads.

In order to analyze market effects of the DR participation in the wholesale energy market as a DR aggregator, in this chapter we adopt the Cournot game model, which is widely used to describe the behavior of traditional generators in the electricity market [68–75]. In contrast to market analysis using the supply function equilibrium (SFE) model [76–78], the Cournot model is generally computationally tractable and much easier to analyze. More detailed reviews of game-theoretic models for the wholesale electricity market can be found in our work [79]. In the Cournot model, each strategic generator chooses its generation level in order to compete in the market and maximize its own profits. The price of electricity is determined by the ISO using these offered generation quantities and a given demand function. A major contribution of our work is to demonstrate how the Cournot game model is modified when a large DR aggregator participates in the game by not just reducing load but queuing and shifting loads in time. Given that load reduction can be compensated as generation, the DR aggregator can now compete with traditional generators in a Cournot-Nash manner to make a profit in the wholesale electricity market.

There are some existing studies that examine inclusion of DR entities in wholesale electricity markets. [80] incorporated the hourly demand response into security-constrained unit commitment (SCUC) for economic and security purposes, where the responsive loads are modeled with their inter-temporal characteristics. [81] presented a stochastic model to schedule reserves as an ancillary service provided by DR in the wholesale electricity market. The proposed stochastic model is formu-

lated as a two-stage stochastic mixed-integer programming (SMIP) problem. [82] proposed a security-constrained forward market clearing algorithm within which the inherent characteristics of demand flexibility are acknowledged when the provision of reserve from the demand side is considered. [83] applied a Cournot game model to examine the specific effect of pumped hydro storage in Germany. The integration of EVs as specific DR resources into wholesale electricity markets is discussed in [84], [85]. Some game-theoretic methods applied to smart grid can be found in [86] and [87].

The remainder of this chapter is organized as follows. Section 6.2 describes the computationally efficient model of deferrable load. Section 6.3 describes the Cournot game model in which DR aggregators participate in the wholesale market from the supply side. Section 6.4 analyzes the equilibrium of the DR-participating Cournot game. Section 6.5 presents numerical results and we conclude our discussion in Section 6.6.

## 6.2 Modeling the Deferrable Load

In [57], a computationally efficient model is proposed to account for the underlying components that constitute the aggregate deferrable load. This model can mainly be used to account for non-interruptible long duration loads with predictable job lengths. Examples include electric vehicles, washing machines, tumble dryers, etc. In the following part, we show how this model for aggregate deferrable load is integrated in the wholesale dispatch optimization.

We assume that electric demand due to any deferrable appliance can be approximated by one of  $q = 1, \dots, Q$  discrete time pulses  $g_q(\cdot)$ . We associate a service

queue with each of these pulses. Every appliance that arrives in the system can be categorized as a deferrable load. It is assigned to one of these queues and needs to wait in order to be authorized to turn on. Denote the total number of arrivals and departures (appliance activations) in the  $q$ -th queue, from the origin of time until epoch  $\ell$ , by  $a_q(\ell)$  and  $d_q(\ell)$  respectively. Note that this means that the number of appliance arrivals and activations at epoch  $\ell$  is respectively given by  $a_q(\ell) - a_q(\ell - 1)$  and  $d_q(\ell) - d_q(\ell - 1)$ . Now, we can write the aggregate deferrable load  $L^S(\ell)$  at time  $\ell$  as,

$$L_S(\ell) = \sum_{q=1}^Q \sum_{k=1}^{\ell} [d_q(k) - d_q(k-1)] g_q(\ell - k). \quad (6.1)$$

Details of this formulation can be found in [57]. In addition, we need to specify the constraints for scheduling the deferrable load as follows.

**(1) Causality and deadline constraints:** From (6.1) we can see that determining the optimal values of  $L^S(\ell)$  is equivalent to finding the optimal values of the  $Q$  queue departure processes  $d_q(\ell)$ . However, several constraints for this departure process  $d_q(\ell)$  need to be specified.

First, we look at simple causality constraints. An appliance cannot be turned on before it requests energy. Also, by definition, the values of  $d_q(\ell)$  should be non-decreasing. Thus,

$$0 \leq d_q(\ell - 1) \leq d_q(\ell) \leq a_q(\ell), \quad \forall \ell, q. \quad (6.2)$$

Additional constraints can be added to ensure that an acceptable quality of service (QoS) is delivered to the deferrable loads. For example, we can ensure that

all the deferrable loads participating in the wholesale market are served by the last hour of the day  $t = H$ . If we denote by  $\tau_q$  the length of the pulse  $g_q(\cdot)$  associated to the  $q$ -th queue, this would mean that:

$$d_q(H - \ell) = a_q(H - \ell), \quad \forall q, \ell < \tau_q. \quad (6.3)$$

However, ensuring that all the loads are served by the end of the day may not result in an acceptable level of QoS and is too optimistic in terms of modeling the flexibility of customers. To penalize the operator for overly delaying the service to the appliances waiting in queues, we propose one of the following options depending on complexity and reliability preferences:

**(2) Per queue maximum delay constraints:** Denote the maximum delay that appliances in the  $q$ -th queue can tolerate by  $\gamma_q$ . Then, maximum delay constraints for the  $q$ -th queue are:

$$d_q(\ell) > a_q(\ell - \gamma_q), \quad \forall q, \ell. \quad (6.4)$$

**(3) Addition of an average delay cost:** A less computationally intensive option is to define a new cost term that can be added to the dispatch optimization in order to avoid dispatches with unacceptable levels of delay. Here, the operator will define a cost term proportional to the average delay experienced by the entire population. The number of appliances waiting in the  $q$ -th queue at time  $\ell$  without receiving service is given by  $a_q(\ell) - d_q(\ell)$ . Thus, if we weigh queues differently, the delay cost increment at time  $\ell$  is given by:

$$\text{DCI}(\ell) = \sum_{q=1}^Q v_q(\ell) [a_q(\ell) - d_q(\ell)]. \quad (6.5)$$

## 6.3 Cournot Game of Wholesale Market with DR Aggregator Participation

We consider an electric network with  $N = |\mathcal{N}|$  nodes, where  $\mathcal{N}$  denotes the set of all nodes in the network. Let  $\mathcal{N}^g \subset \mathcal{N}$  denote nodes at which traditional generators are located. DR aggregators are located at a set of nodes, which we denote by  $\mathcal{N}^G$ . In addition, we assume  $\mathcal{N}^g \cap \mathcal{N}^G = \emptyset$ .

We model the day-ahead energy market as Cournot game, which has been well established in a series of papers [70], [74], [73] and [75] by Oren et al. In this model, suppliers determine generation quantities to compete in a Nash-Cournot manner over a time horizon  $H$ . The suppliers here include both traditional generators and DR aggregators. Each traditional generator  $i \in \mathcal{N}^g$  decides its generation quantities as  $\zeta_i^g(\ell)$  for  $\ell = 1, 2, \dots, H$ , while the scheduling capability enables DR aggregator to make a load modification decision  $\zeta_j^{dl}(\ell)$ , for  $j \in \mathcal{N}^{dl}$ ,  $\ell = 1, 2, \dots, H$ . Both traditional generators and the DR aggregator aim to maximize their profit, given as the revenue collected minus the generation cost.

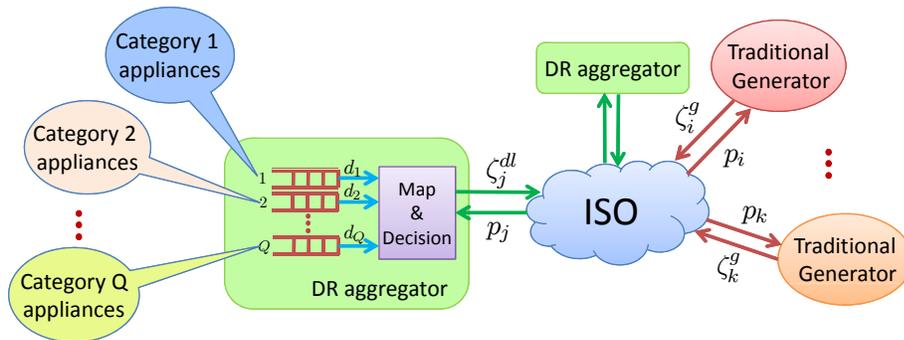


Figure 6.1: Participation of DR aggregators in the wholesale market

The ISO in this model determines how to dispatch electricity within the network so as to maximize total social surplus subject to transmission constraints. Specif-

ically, it obtains the optimal power import value  $r_i(\ell)$  at each node. Since the generator and the ISO take into consideration each other's actions, their decisions are treated as simultaneous moves in the model. The structure of the wholesale market with inclusion of DR aggregators is shown in Figure 6.1.

In Cournot model, the demand behavior at node  $i$  is described by the inverse demand function (IDF) [74] , [70], and is given as  $p = P_i(w)$ , where  $p$  is the price of electricity and  $w$  is the demand. This function describes how the demand responds to the price, and in general it is a decreasing function indicating that as price increases, customers will decrease their demand. Given IDF  $P_i(\cdot)$  at each node  $i$ , the demand as well as the locational price for each node can be obtained by jointly solving the optimization problem of generators, DR aggregators and ISO (i.e., the equilibrium point). We will formulate optimization problems for each entity in next section.

For ease of analysis, we employ the linear demand function, which is widely used in the Cournot game model in Oren's work [70], [74], [73] and [75], as

$$p_j = P_j(w_j) = c_j - b_j \cdot w_j, \quad j \in \mathcal{N} \quad (6.6)$$

where  $w_j$  is the demand at node  $j$ . The slope  $b_j$  indicates the demand responsiveness, i.e., how the price affects the quantity demanded, while  $c_j$  indicates factors other than price that affect demand. The parameter  $b_j$  could be time-varying, as  $b_j(\ell)$  for hour  $\ell$ .

### 6.3.1 Optimization for generators

Traditional generators aim to maximize their profit by simultaneously choosing generation quantities  $\zeta_i^g(\ell)$ ,  $i \in \mathcal{N}^g$ ,  $\ell = 1, 2, \dots, H$ . Since there is no temporal

correlation of each hour's generation,<sup>1</sup> the profit maximization for the whole day is just to maximize the profit for each hour independently. For notational simplicity, we omit the hour index  $\ell$  here. The profit maximization problem for a generator located at node  $i$  can be formulated as:

$$\max_{\zeta_i^g: i \in \mathcal{N}^g} P_i(r_i + \zeta_i^g)\zeta_i^g - C_i(\zeta_i^g) \quad (6.7)$$

$$\text{s.t. } \zeta_i^{\min} \leq \zeta_i^g \leq \zeta_i^{\max}, \quad i \in \mathcal{N}^g \quad (6.8)$$

where  $P_i(r_i + \zeta_i^g)$  is the price with demand of  $r_i + \zeta_i^g$ . Note that  $r_i$  is the power import value at node  $i$  determined by ISO. When  $r_i > 0$ , power is delivered to this node; when  $r_i$  is negative, node  $i$  transmits power into the electricity network. The first term in the objective function (6.7) represents the revenue of generating  $\zeta_i^g$  of power, while the second term refers to the generation cost. The cost of generation is usually modeled to be convex, e.g., quadratic or piece-wise linear function. The  $\zeta_i^{\min}$  and  $\zeta_i^{\max}$  in the constraint indicate the minimum generation when ON and capacity limit of the generator located at node  $i$ . The generators here are assumed to have been scheduled ON in the unit commitment (UC) problem.

### 6.3.2 Optimization for DR aggregator

With schedulability from deferrable loads, the DR aggregator can make load reductions, which is equivalent to generation in terms of balancing power on the electricity market. In order to formulate the optimization problem for the DR aggregator, how the 'generation' value  $\zeta^{dl}(\ell)$  are mapped in terms of deferrable load parameters (e.g., the departure process and the  $q$ -th queue's load profile) as well as constraints of these parameters need to be modeled mathematically. The

---

<sup>1</sup>We omit the ramp constraint for simplicity.

load modification  $\zeta^{dl}(\ell)$  from deferrable load managed by a DR aggregator can be computed as

$$\zeta^{dl}(\ell) = \bar{L}_S(\ell) - L_S(\ell) = \sum_{q=1}^Q \sum_{k=1}^{\ell} [\lambda_q(k) - d_q(k) + d_q(k-1)] g_q(\ell - k), \quad (6.9)$$

where  $\bar{L}_S(\ell) = \sum_{q=1}^Q \sum_{k=1}^{\ell} \lambda_q(k) g(\ell - k)$  is the original demand from these deferrable loads, and  $\lambda_q(k) = [a_q(k) - a_q(k-1)]$  is the arrival rate for time  $k$ . As discussed in Section 6.2, the DCI can be viewed as cost that customers experience due to a ‘generation’ of  $\zeta^{dl}(\ell)$ .

By choosing the departure processes  $\{d_{q,j}(\ell)\}$ , the DR aggregator aims to maximize its profit over the time horizon  $H$ , which is the money it gets for load modification  $\zeta_j^{dl}$ , minus the inconvenience cost DCI. For notational simplicity, we omit the DR aggregator index subscript of  $j$ , and the optimization for the DR aggregator is:

$$\max_{\{d_q(\ell)\}} \sum_{\ell=1}^H [P_{\ell} (r(\ell) + \zeta^{dl}(\ell)) \zeta^{dl}(\ell)] - \text{DCI} \quad (6.10)$$

$$\text{s.t. } d_q(H - \ell) = a_q(H - \ell), \quad \forall q, \ell < \tau_q \quad (6.11)$$

$$d_q(\ell) \leq a_q(\ell) = \sum_{s=1}^{\ell} \lambda_q(s), \quad \forall q, \ell \quad (6.12)$$

$$d_q(\ell - 1) \leq d_q(\ell), \quad \forall q, \ell \quad (6.13)$$

$$d_q(\ell) \geq a_q(\ell - \gamma_q), \quad \forall q, \ell \quad (6.14)$$

$$d_q(\ell) \in \mathbb{Z}_+ \cup \{0\}, \quad \forall q, \ell \quad (6.15)$$

where  $\zeta^{dl}(\ell)$  and DCI are expressed as in (6.9) and (6.5), respectively. The constraints (6.11)-(6.15) are corresponding to (6.2), (6.3) and (6.4) in Section 6.2.

### 6.3.3 Optimization for the ISO

The ISO aims to maximize the social welfare of the entire system, which is given by the total consumer willingness-to-pay (the aggregated area under the nodal inverse demand curves) minus the total generation cost. (See [74], [88] for details.) The social welfare of node  $i$  can be written as:  $\sum_{i \in \mathcal{N}} \left( \int_0^{r_i(\ell) + \zeta_i(\ell)} P_i(\tau_i) d\tau_i - C_i(\zeta_i(\ell)) \right)$ , where  $\zeta_i(\ell)$  is the  $\ell$ -th hour's generation value ( $\zeta_i^g(\ell)$  if  $i \in \mathcal{N}^g$ , or  $\zeta_i^{dl}(\ell)$  if  $i \in \mathcal{N}^{dl}$ ) at node  $i$ .  $C_i(\cdot)$  is the corresponding generation cost function (delay inconvenience cost) for generators (DR aggregator). The  $\{r_i(\ell)\}_{i \in \mathcal{N}}$  is the power import value for each node, which is decided by the ISO. We next specify the constraints for ISO optimization.

**(1) Power balance constraint:** Since  $r_i(\ell)$ ,  $i \in \mathcal{N}$  is the power import value at node  $i$ , the demand at node  $i$  is actually  $\zeta_i^g(\ell) + r_i(\ell)$ , if there is a traditional generator, and  $\zeta_i^{dl}(\ell) + r_i(\ell)$ , if a DR aggregator is located at node  $i$ . Since the total load and total generation must be balanced at all moments,  $\{r_i(\ell)\}$  should satisfy the energy balancing constraint:  $\sum_{i \in \mathcal{N}} r_i(\ell) = 0$ .  $\ell = 1, 2, \dots, H$

**(2) Demand constraint:** For all nodes, the demand value should not be negative, i.e.,  $\zeta_i + r_i \geq 0$ ,  $i \in \mathcal{N}$ . For the DR aggregator node, although it has an equivalent 'generation'  $\zeta^{dl}(\ell)$ , it is actually a demand node, i.e., it consumes power from the network. So the power import at a DR aggregator node should be a non-negative value, i.e.,  $r_i(\ell) \geq 0$ ,  $i \in \mathcal{N}^{dl}$ .

**(3) Transmission constraint:**  $\{r_i(\ell)\}$  must satisfy the network line power constraint, i.e., the resulting power flows should not exceed the thermal limits of the

corresponding transmission line in both directions. We use Power Transfer Distribution Factor (PTDF) to describe the linear approximation of the relation between power flow and power injection. Given the PTDF of node  $i$  in relation to branch  $m$  as  $S_{mi}$ , if the injection change at node  $i$  is  $\Delta r_i$ , the power flow change in the branch  $m$  would be  $\Delta P_m = \sum_{i=1}^N S_{mi} \Delta r_i$ . Let  $\mathcal{M}$  denote the set of all the transmission lines in the system and  $K_m$  denote the limit of transmission line  $m$ . The transmission constraints in terms of  $\{r_i\}$  is:  $-K_m \leq \sum_{i \in \mathcal{N}} S_{mi} r_i \leq K_m$ ,  $m \in \mathcal{M}$ .

Note that the  $\zeta_i(\ell)$  values, decided by generators/DR aggregator, are known to the ISO, so the cost function part can be removed from the objective function. Hence, the optimization for ISO for hour  $\ell$  can be formulated as

$$\max_{\{r_i\}} \sum_{i \in \mathcal{N}} \int_0^{r_i(\ell) + \zeta_i(\ell)} P_{i,\ell}(\tau_i) d\tau_i \quad (6.16)$$

$$\text{s.t.} \quad \sum_{i \in \mathcal{N}} r_i(\ell) = 0 \quad (6.17)$$

$$r_i(\ell) + \zeta_i(\ell) \geq 0, \quad i \in \mathcal{N} \quad (6.18)$$

$$r_i(\ell) \geq 0, \quad i \in \mathcal{N}^{dl} \quad (6.19)$$

$$-K_m \leq \sum_{i \in \mathcal{N}} S_{mi} r_i(\ell) \leq K_m, \quad m \in \mathcal{M}, \quad (6.20)$$

## 6.4 Equilibrium Analysis

The equilibrium conditions are the result of jointly solving the optimization problems for the ISO, generators and DR aggregators. It is easy to see that the optimization for each entity is convex so the optimal (KKT) conditions can be derived. In the following parts, we let bold font  $\mathbf{X}$  denote matrix,  $\vec{X}$  denote vector,  $(\mathbf{X})_{i,j}$  denote  $i, j$  entry of matrix  $\mathbf{X}$ ,  $\vec{X}^T(\mathbf{X}^T)$  denote transpose of vector  $\vec{X}$  (matrix  $\mathbf{X}$ ),

and  $\text{diag}\{\vec{x}\}$  denote the diagonal matrix with diagonal entries as vector  $\vec{x}$ .

### 6.4.1 Optimality conditions for DR aggregator

Let vector  $\vec{D} = [\vec{d}_1, \vec{d}_2, \dots, \vec{d}_Q]^T$  denote decision variables of the DR aggregator, where each element is a row vector  $\vec{d}_q = [d_q(1), d_q(2), \dots, d_q(H)]$ . To express the optimization in a general matrix form, we denote arrival rate vector by  $\vec{\Lambda} = [\vec{\lambda}_1, \vec{\lambda}_2, \dots, \vec{\lambda}_Q]^T$ , where  $\vec{\lambda}_q = [\lambda_q(1), \lambda_q(2), \dots, \lambda_q(H)]$ ; load profile matrix by  $\mathbf{G} = [\vec{G}(1), \vec{G}(2), \dots, \vec{G}(H)]^T$ , where each row is a transpose of a vector  $\vec{G}(\ell)$  with length  $H \cdot Q$  given as:  $\vec{G}(\ell) = [\vec{g}_1(\ell), \vec{g}_2(\ell), \dots, \vec{g}_Q(\ell)]^T$ , with  $\vec{g}_q(\ell) = [g_q(\ell - 1), g_q(\ell - 2), \dots, g_q(\ell - H)]$ , and we define  $g_q(\ell - k) = 0$  if  $\ell - k < 0$ .

Define matrix  $\mathbf{U}$  as

$$\mathbf{U} = \mathbf{I}_Q \otimes \tilde{\mathbf{U}}, \quad (\tilde{\mathbf{U}})_{i,j} = \begin{cases} -1, & i = j \\ 1, & i - 1 = j, i, j \in [1, H] \\ 0, & \text{o.w.} \end{cases}$$

where  $\mathbf{I}_Q$  is identity matrix of size  $Q$ , and  $\otimes$  denotes the Kronecker product. So  $\zeta^{dl}(\ell)$  can be expressed in matrix form as

$$\zeta^{dl}(\ell) = \vec{G}(\ell)^T (\vec{\Lambda} + \mathbf{U}\vec{D}).$$

The arrival process vector of appliances ( $\vec{A}$ ) can be expressed in terms of arrival

rate vector  $\vec{\Lambda}$  as  $\vec{A} = \mathbf{L}\vec{\Lambda}$ , where  $\mathbf{L}$  is defined as

$$\mathbf{L} = \mathbf{I}_Q \otimes \tilde{\mathbf{L}}, \quad (\tilde{\mathbf{L}})_{i,j} = \begin{cases} 1, & i \geq j \\ 0, & \text{o.w.} \end{cases}, i, j \in [1, H]$$

The delay inconvenience cost (DCI) in (6.5) can also be expressed in matrix form as

$$\text{DCI} = \vec{v}^T(\mathbf{L}\vec{\Lambda} - \vec{D}),$$

where  $\vec{v} = [\vec{v}_1^D, \vec{v}_2^D, \dots, \vec{v}_Q^D]^T$  is vector of delay cost per hour, such that  $\vec{v}_q^D = [v_q^D(1), v_q^D(2), \dots, v_q^D(H)]$ . Thus,  $\vec{v}$  has a length of  $H \cdot Q$ . The inequality constraint (6.12) and (6.13) can be expressed as

$$\vec{D} - \mathbf{L}\vec{\Lambda} \leq \vec{0}, \tag{6.21}$$

$$\mathbf{U}\vec{D} \leq \vec{0}, \tag{6.22}$$

and the inequality constraint (6.14) can be expressed as

$$\vec{D} \geq \mathbf{N}\mathbf{L}\vec{\Lambda}, \tag{6.23}$$

where  $\mathbf{N}$  is defined as

$$\mathbf{N} = \begin{bmatrix} \tilde{\mathbf{N}}_1 & 0 & \cdots & 0 \\ 0 & \tilde{\mathbf{N}}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\mathbf{N}}_Q \end{bmatrix}, \quad (\tilde{\mathbf{N}}_q)_{i,j} = \begin{cases} 1, & \forall i \in [1, H], j = i - \tau_q > 0, \\ 0, & \text{o.w.} \end{cases}$$

So we can see  $\mathbf{N}$  is a square matrix with size of  $H \cdot Q$ .

The equality constraint (6.11) can be expressed as

$$\mathbf{M}\vec{D} = \mathbf{M}\mathbf{L}\vec{\Lambda}, \quad (6.24)$$

where the matrix  $\mathbf{M}$  is defined as

$$\mathbf{M} = \begin{bmatrix} \tilde{\mathbf{M}}_1 & 0 & \cdots & 0 \\ 0 & \tilde{\mathbf{M}}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\mathbf{M}}_Q \end{bmatrix}, \quad (\tilde{\mathbf{M}}_q)_{i,j} = \begin{cases} 1, & \forall i \in [1, \gamma_q], j = H - i + 1 \\ 0, & \text{o.w.} \end{cases}$$

So we can see  $\mathbf{M}$  has rows of  $\sum_{q=1}^Q \gamma_q$  elements and columns of  $H \cdot Q$  elements.

With linear IDF defined in (6.6), the objective function of DR aggregator can be written as  $\sum_{\ell=1}^H [c(\ell)\zeta^{dl}(\ell) - b(\ell)r(\ell)\zeta^{dl}(\ell) - b(\ell)(\zeta^{dl}(\ell))^2] - \text{DCI}$ . Note that we omit the subscript  $j$  for simplicity. The matrix form of the objective function can then be written as

$$z + \vec{Y}^T \vec{D} + \vec{D}^T \mathbf{X} \vec{D}. \quad (6.25)$$

where constant  $z$ , vector  $\vec{Y}$ , and matrix  $\mathbf{X}$  are

$$\begin{aligned} z &= \vec{c}^T \mathbf{G} \vec{\Lambda} - \vec{r}^T \mathbf{B} \mathbf{G} \vec{\Lambda} - \vec{v}^T \mathbf{L} \vec{\Lambda} - \vec{\Lambda}^T \mathbf{G}^T \mathbf{B} \mathbf{G} \vec{\Lambda}, \\ \vec{Y}^T &= (\vec{c}^T - \vec{b}^T \mathbf{R} - 2\vec{\Lambda}^T \mathbf{G}^T \mathbf{B}) \mathbf{G} \mathbf{U} + \vec{v}^T, \\ \mathbf{X} &= -\mathbf{U}^T \mathbf{G}^T \mathbf{B} \mathbf{G} \mathbf{U}, \end{aligned}$$

such that  $\vec{c} = [c(1), c(2), \dots, c(H)]^T$ ,  $\vec{b} = [b(1), b(2), \dots, b(H)]^T$  with  $\mathbf{B} = \text{diag}\{\vec{b}\}$  and  $\vec{r} = [r(1), r(2), \dots, r(H)]^T$ . Since  $z$  is a constant, it can be removed from the objective function.

Note that since  $d_q(\ell)$  must take non-negative integer values, this optimization is an integer programming problem, so the optimality conditions can not be derived directly. Here we relax  $d_q(\ell)$  to be a continuous variable, and we show that the floor value of the solution for the relaxed problem is feasible, as shown in Lemma 6.4.1.

**Lemma 6.4.1.** *If the optimal solution to the relaxation problem, in which  $d_q(\ell)$  is relaxed as a continuous non-negative value, is  $d_q^*(\ell)$ , then the floor of this value, i.e.,  $\lfloor d_q^*(\ell) \rfloor$  is always feasible.*

*Proof.* Let  $\vec{D}^*$  denote the vector of optimal solutions of departure process  $\{d_q^*(\ell)\}$  to the relaxation problem. Since  $\lfloor x \rfloor \leq x$ , we get

$$\lfloor \vec{D}^* \rfloor \leq \vec{D}^* \leq \mathbf{L}\vec{\Lambda}.$$

So the solution  $\lfloor \vec{D}^* \rfloor$  satisfies constraint (6.21). Also since  $d_q^*(\ell - 1) \leq d_q^*(\ell)$ , the floor of  $d_q^*(\ell)$  satisfies  $\lfloor d_q^*(\ell - 1) \rfloor \leq \lfloor d_q^*(\ell) \rfloor$ . Thus, the constraint (6.22) is also satisfied. For inequality constraint (6.23), since  $\mathbf{NL}\vec{\Lambda}$  is an integer vector, we get  $\lfloor \vec{D}^* \rfloor \geq \mathbf{NL}\vec{\Lambda}$ .

For the equality constraints, the original solutions  $\{d_q^*(H - \ell)\}$ ,  $\ell < \tau_q$ , are integer values, so  $\lfloor d_q^*(H - \ell) \rfloor = d_q^*(H - \ell)$ . This means that  $\lfloor \vec{D}^* \rfloor$  satisfies constraint (6.24).

So the floor solution  $\lfloor \vec{D}^* \rfloor$  is feasible. □

Using Lemma 6.4.1, we can get the optimality conditions of the relaxed problem and floor it to get the (sub-optimal) final solution. The relaxed problem can be

written as a standard quadratic program as

$$\min_{\vec{D}} f(\vec{D}) = -\vec{Y}^T \vec{D} - \vec{D}^T \mathbf{X} \vec{D} \quad (6.26)$$

$$\text{s.t. } \mathbf{M} \vec{D} - \mathbf{M} \mathbf{L} \vec{\Lambda} = \vec{0} \quad (\vec{\Theta}_1) \quad (6.27)$$

$$\mathbf{U} \vec{D} \leq \vec{0} \quad (\vec{\Theta}_2) \quad (6.28)$$

$$-\vec{D} + \mathbf{N} \mathbf{L} \vec{\Lambda} \leq \vec{0} \quad (\vec{\Theta}_3) \quad (6.29)$$

$$\vec{D} - \mathbf{L} \vec{\Lambda} \leq \vec{0}, \quad (\vec{\Theta}_4) \quad (6.30)$$

where  $\vec{\Theta}_1$ – $\vec{\Theta}_4$  are the corresponding Lagrangian multiplier vectors. The KKT conditions can be written accordingly:

$$-\vec{Y} - 2\mathbf{X} \vec{D} + \mathbf{M}^T \vec{\Theta}_1 + \mathbf{U}^T \vec{\Theta}_2 - \vec{\Theta}_3 + \vec{\Theta}_4 = \mathbf{0} \quad (6.31)$$

$$\mathbf{M} \vec{D} - \mathbf{M} \mathbf{L} \vec{\Lambda} = \vec{0} \quad (6.32)$$

$$\mathbf{U} \vec{D} \leq \vec{0} \quad (6.33)$$

$$-\vec{D} + \mathbf{N} \mathbf{L} \vec{\Lambda} \leq \vec{0} \quad (6.34)$$

$$\vec{D} - \mathbf{L} \vec{\Lambda} \leq \vec{0} \quad (6.35)$$

$$\vec{\Theta}_2, \vec{\Theta}_3, \vec{\Theta}_4 \geq \vec{0} \quad (6.36)$$

$$\text{diag}\{\vec{\Theta}_2\} \mathbf{U} \vec{D} = \vec{0} \quad (6.37)$$

$$\text{diag}\{\vec{\Theta}_3\} (\vec{D} - \mathbf{N} \mathbf{L} \vec{\Lambda}) = \vec{0} \quad (6.38)$$

$$\text{diag}\{\vec{\Theta}_4\} (\vec{D} - \mathbf{L} \vec{\Lambda}) = \vec{0} \quad (6.39)$$

## 6.4.2 Optimality conditions for the ISO and generators

Since the optimization variables of ISO and generators are not time correlated, they can be conducted independently for each hour. For ease of notational simpli-

fication, we omit the time index  $\ell$ . For the ISO, the power import for each node can be expressed as a vector  $\vec{R} = [r_1, r_2, \dots, r_N]^T$ . Let  $\vec{b}_n = [b_1, b_2, \dots, b_N]^T$  and  $\vec{c}_n = [c_1, c_2, \dots, c_N]^T$  denote the slope and intercept vector of linear IDF in (6.6) for all nodes, respectively, and  $\mathbf{B}_n = \text{diag}\{\vec{b}_n\}$ . Let vector  $\vec{\zeta} = [\zeta_1, \zeta_2, \dots, \zeta_N]^T$  represent  $\zeta_i$  for all nodes, where  $\zeta_i$  can be represented as follows.

$$\zeta_i = \begin{cases} \zeta_i^g, & \text{if } i \in \mathcal{N}^g, \\ \zeta_i^{dl}, & \text{if } i \in \mathcal{N}^{dl}, \\ 0, & \text{otherwise.} \end{cases} \quad (6.40)$$

The optimization (6.16)-(6.20) can be written in matrix forms as

$$\min_{\vec{R}} (\vec{\zeta}^T \mathbf{B}_n - \vec{c}_n^T) \vec{R} + \frac{1}{2} \vec{R}^T \mathbf{B}_n \vec{R} \quad (6.41)$$

$$\text{s.t. } \vec{1}^T \vec{R} = 0 \quad (\Phi_1) \quad (6.42)$$

$$-\vec{R} - \vec{\zeta} \leq \vec{0} \quad (\vec{\Phi}_2) \quad (6.43)$$

$$-\mathbf{I}^{dl} \vec{R} \leq \vec{0} \quad (\vec{\Phi}_3) \quad (6.44)$$

$$-\mathbf{S} \vec{R} - \vec{K} \leq \vec{0} \quad (\vec{\Phi}_4) \quad (6.45)$$

$$\mathbf{S} \vec{R} - \vec{K} \leq \vec{0} \quad (\vec{\Phi}_5) \quad (6.46)$$

where  $\vec{1} = [1, 1, \dots, 1]^T$  with length of  $N$ , and diagonal matrix  $\mathbf{I}^{dl} = \text{diag}\{\vec{1}_{\mathcal{N}^{dl}}\}$ , with  $\vec{1}_{\mathcal{N}^{dl}}$  denoting the unit vector representing the DR aggregator nodes.  $\mathbf{S}$  is the PTDF matrix with size of  $M \times N$ . Let  $\Phi_1, \vec{\Phi}_2$ - $\vec{\Phi}_5$  denote the Lagrange multiplier (vectors) of equality and inequality constraints, and the KKT conditions can be written accordingly.

$$(\vec{c}_n - \mathbf{B}_n \vec{\zeta}) - \mathbf{B}_n \vec{R} + \Phi_1 \vec{1} - \vec{\Phi}_2 - \mathbf{I}_G \vec{\Phi}_3 - \mathbf{S}^T \vec{\Phi}_4 + \mathbf{S}^T \vec{\Phi}_5 = \vec{0} \quad (6.47)$$

$$\vec{1}^T \vec{R} = 0 \quad (6.48)$$

$$-\vec{R} - \vec{\zeta} \leq \vec{0} \quad (6.49)$$

$$-\mathbf{I}^{dl} \vec{R} \leq \vec{0} \quad (6.50)$$

$$-\mathbf{S} \vec{R} - \vec{K} \leq \vec{0} \quad (6.51)$$

$$\mathbf{S} \vec{R} - \vec{K} \leq \vec{0} \quad (6.52)$$

$$\vec{\Phi}_2, \vec{\Phi}_3, \vec{\Phi}_4, \vec{\Phi}_5 \geq \vec{0} \quad (6.53)$$

$$\text{diag}\{\vec{\Phi}_2\}(\vec{R} + \vec{\zeta}) = \vec{0} \quad (6.54)$$

$$\text{diag}\{\vec{\Phi}_3\} \mathbf{I}^{dl} \vec{R} = \vec{0} \quad (6.55)$$

$$\text{diag}\{\vec{\Phi}_4\}(\mathbf{S} \vec{R} + \vec{K}) = \vec{0} \quad (6.56)$$

$$\text{diag}\{\vec{\Phi}_5\}(\mathbf{S} \vec{R} - \vec{K}) = \vec{0} \quad (6.57)$$

For the generator at node  $i$ , let  $\xi_i^-$  and  $\xi_i^+$  denote the Lagrange multiplier of inequality constraints for lower and upper bound of  $\zeta_i^g$ , respectively. The KKT conditions can be formulated accordingly.

$$-c_i + b_i r_i + 2b_i \zeta_i^g + C'_i(\zeta_i^g) - \xi_i^- + \xi_i^+ = 0 \quad (6.58)$$

$$\xi_i^- \geq 0, \quad \xi_i^+ \geq 0 \quad (6.59)$$

$$\zeta_i^{\min} \leq \zeta_i^g \leq \zeta_i^{\max} \quad (6.60)$$

$$\xi_i^- \cdot (\zeta_i^{\min} - \zeta_i^g) = 0 \quad (6.61)$$

$$\xi_i^+ \cdot (\zeta_i^g - \zeta_i^{\max}) = 0, \quad (6.62)$$

where  $C'_i(\zeta_i^g)$  is the first derivative of the generator cost function.

### 6.4.3 Existence and uniqueness of the equilibrium

The equilibrium point can be solved by jointly solving the equilibrium conditions for DR aggregators, generators, and the ISO. With linear inverse demand function, and cost function of generators being convex quadratic or piece-wise linear, which is a common assumption, we have Lemma 6.4.2 about the existence and uniqueness of the equilibrium.

**Lemma 6.4.2.** *If we assume a convex cost function, e.g., quadratic or piece-wise linear, for generators, and linear inverse demand function, the equilibrium point of the Cournot game with participating DR aggregators exists and is unique.*

*Proof.* a) For the DR aggregator's optimization in (6.26), since  $-\mathbf{X} = \mathbf{U}^T \mathbf{G}^T \mathbf{B} \mathbf{G} \mathbf{U}$  is a positive definite matrix ( $\mathbf{B}$  is positive definite diagonal matrix), the objective is a convex function. So the utility function that the DR aggregator wants to maximize is concave. b) For the ISO's optimization, from (6.16) we can see the objective is also a convex function, implying the utility function is concave. c) For generators' optimization in (6.7), if the cost function  $C_i(\zeta_i^g)$  is a convex (quadratic or piece-wise linear as general forms), the utility function is a concave function. d) The constraints for these optimizations are linear. In this case, the existence of an equilibrium directly results from Theorem 1 in [89]. Moreover, the equilibrium is unique due to Theorem 2 and 3 in [89].  $\square$

With Lemma 6.4.2, we can get the market equilibrium by jointly solving the optimizations of the DR aggregator, generators, and the ISO.

## 6.5 Numerical Results

We apply the proposed Cournot model to the IEEE 24-bus test network shown in Figure 6.2, which is composed of 24 nodes and 38 lines. The time horizon is set as  $H = 24$  hours. We assume the generators' cost function as quadratic, where the coefficients of cost functions, lower and upper bounds of generation, as well as the transmission line parameters for this 24-bus network can be found in case file 'case24\_ieee\_rts.m' in MATPOWER [90]. The function 'makePTDF' in MATPOWER can be used to compute the PTDF parameters for the system. We choose nodes 20, 19, 17, and 6 as candidates for DR aggregators shown in Figure 6.2. The DR aggregator is responsible for providing battery recharging services to a population of 40k plug-in hybrid electric vehicles (PHEVs), where a realistic arrival, charge request, and maximum delay model developed in [91] is employed in the simulation. We assume all nodes have the same IDF intercept

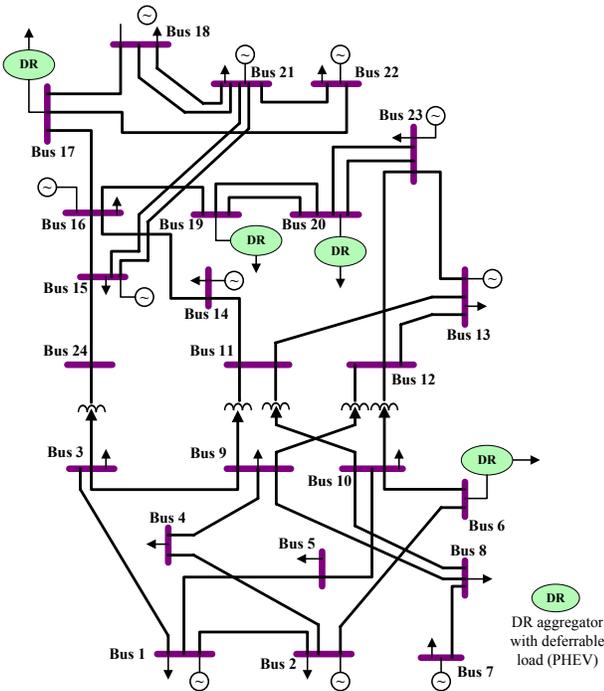


Figure 6.2: IEEE 24-bus test network with DR aggregators

$c = \$70$ , which keeps unchanged over time  $\ell$ . The IDF slope  $b_i$  at each node employs the values in [73]. The CPLEX 12.4 [39] optimization solver is used to solve the quadratic programming problem for DR aggregators, generators, and the ISO. MATLAB 2009a is used to formulate the problem and link the CPLEX solver. The simulation environment is of Intel Duo Core 2.0GHz with 2GB memory. The execution time for solving the DR aggregator’s optimization in (6.26) with  $H = 24$  hours and the number of queues  $Q = 5$  is 0.016s.

In order to make this Cournot game energy market similar to the real market, we need to adjust the slope of each hour in a day such that the hours with a higher actual price have a lower slope; thus gives a higher price in Cournot model. We use day-ahead prices of New York City from NYISO [36] for these adjustments, i.e., given day-ahead LMP as  $\kappa(\ell)$  for hour  $\ell$ , the adjusted coefficient is computed as

$$\beta(\ell) = \frac{(1/24) \sum_{k=1}^{24} \kappa(k)}{\kappa(\ell)},$$

and adjusted time-variant slope at bus  $i$  can be computed as  $b_i(\ell) = \beta(\ell) \cdot b_i$ .

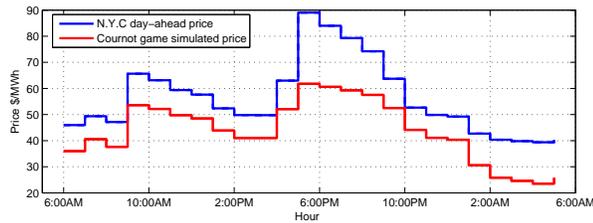


Figure 6.3: Cournot game price validation at node 20 (No DR aggregator)

Figure 6.3 plots the market prices of electricity at node 20 for the 24 hours obtained from the Cournot model, compared with the actual day-ahead price of N.Y.C.; in this case, we do not include the DR aggregator’s participation. We can see the price profile has the same trend as the actual case, verifying our Cournot

game model is reasonable. In the Cournot model, since each node has an inverse demand function, this demand responsiveness serves as a feedback scheme which prevents the price from reaching a very high value. Note that we use a 24-hour planning horizon from 6:00AM to 5:59AM the next day. With this shifting, there is enough off-peak time for the delayed load coming from the peak hours of yesterday to spread, thus avoiding a load bump in the last hours. This time-shift actually does not affect the ISO’s day-ahead optimization since we ignore the generation ramp in our model and there is no time correlation in the optimization.

Figure 6.4 provides the electricity price profile for the day at DR aggregator node (Bus 20) and regular node (Bus 5), using the Cournot model with presence and absence of the DR aggregator. We can see during the original peak price hour

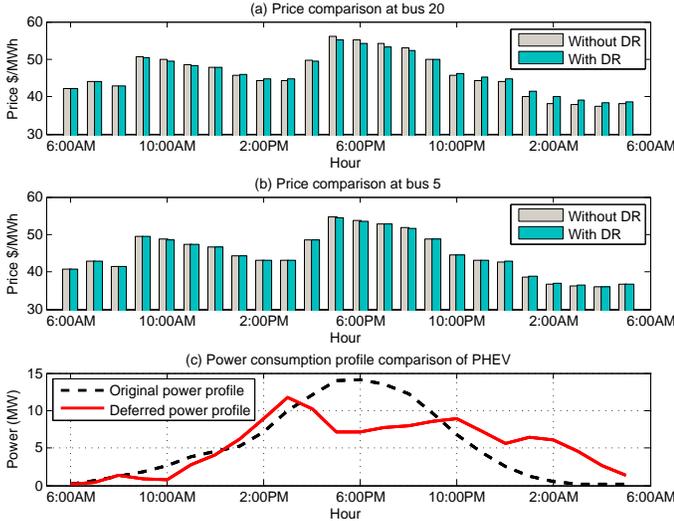


Figure 6.4: Price comparison with/without DR aggregator

(5:00 PM–8:00 PM), the price when DR aggregator participates in the market decreases compared to the case when there is no DR aggregator for both nodes. During peak times, the DR aggregator will benefit more from load reduction due to higher prices. The load reduction will allow other generators to lower their generation, so avoiding the high cost generation and alleviating the congestion make

electricity prices decrease. The price reduction at the DR aggregator node (Bus 20) is more obvious than that at regular nodes. We also notice that during off-peak price hour (12:00 AM – 6:00 AM), the price increases since the DR aggregator shifts deferrable loads during this period, which can be seen in the power consumption profile of aggregated PHEVs load in Figure 6.4(c).

Figure 6.5(a) shows as more load nodes are equipped with DR aggregators, prices decrease further during peak hours. Figure 6.5(b) shows as more DR aggregators are added, the overall profit of DR aggregators increases, i.e., all DR aggregators make profit for participation in the market, which is the customers' saving to enroll in this program.

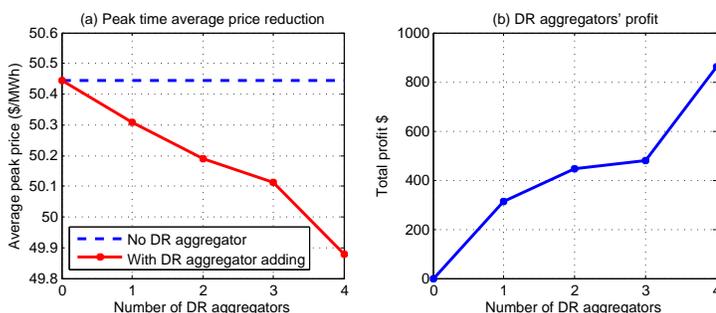


Figure 6.5: Price reduction and profit of DR aggregator

## 6.6 Chapter Summary

In this chapter, we apply a Cournot game model to analyze market effects of DR aggregators' participation in the wholesale market. We show how the deferrable load as a market asset is managed by the DR aggregator to compete with other generators in a Cournot-Nash manner. We formulate optimization problems for DR aggregators, generators and the ISO, and analyze the market equilibrium. Numerical results show that with DR participation, the market price will be lowered and thus save money for customers.

# Bibliography

- [1] U.S. Energy Information Administration (EIA), “Annual Energy Outlook 2011 with Projection to 2035,” 2010. [Online]. Available: <http://www.eia.gov/oiaf/aeo/electricity.html>
- [2] U.S. Department of Energy (DOE), “20% Wind Energy by 2030: Increasing Wind Energy’s Contribution to U.S. Electricity Supply,” DOE Office of Energy Efficiency and Renewable Energy Report, July 2008. [Online]. Available: <http://www1.eere.energy.gov/wind/pdfs/41869.pdf>
- [3] U.S. Department of Energy, “The Smart Grids: An Introduction,” 2009.
- [4] —, “Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them,” 2006. [Online]. Available: <http://eetd.lbl.gov/ea/ems/reports/congress-1252d.pdf>
- [5] CNT Energy and Ameren Illinois Utilities, “Ameren Power Smart Pricing,” 2012. [Online]. Available: <http://www.powersmartpricing.org/how-it-works/>
- [6] CNT Energy and ComEd Utilities, “ComEd Residential Real-Time Pricing (RRTP),” 2012. [Online]. Available: <http://www.cntenergy.org/pricing/comed-rrtp/>

- [7] Wisconsin Public Service, “Contracted Direct Load Control (CDLC),” 2012. [Online]. Available: <http://www.wisconsinpublicservice.com/business/cdlc.aspx>
- [8] conEdison, “Demand Response/Distribution Load Relief Program,” 2012. [Online]. Available: [http://www.coned.com/energyefficiency/dist\\_load\\_relief.asp](http://www.coned.com/energyefficiency/dist_load_relief.asp)
- [9] S. Kishore and L. Snyder, “Control Mechanisms for Residential Electricity Demand in Smart Grids,” in *Proc. of IEEE Conf. on Smart Grid Communications*, 2010.
- [10] F. Rahimi and A. Ipakchi, “Demand Response as a Market Resource Under the Smart Grid Paradigm,” *IEEE Trans. on Smart Grid*, vol. 1, no. 1, pp. 82–88, 2010.
- [11] New York Independent System Operator, “Emergency Demand Response Program Manual,” *Technical Report*, December, 2010.
- [12] PJM Interconnection, “Manual 11: Energy and Ancillary Services Market Operations,” January 2012. [Online]. Available: <http://pjm.com/~media/documents/manuals/m11.ashx>
- [13] ISO New England, “Load Response Program.” [Online]. Available: [http://www.iso-ne.com/regulatory/tariff/sect\\_3/mr1\\_append-e.pdf](http://www.iso-ne.com/regulatory/tariff/sect_3/mr1_append-e.pdf)
- [14] Federal Energy Regulatory Commission (FERC), “Order No. 719, Wholesale Competition in Regions with Organized Electric Markets,” October 2008. [Online]. Available: <http://www.ferc.gov/whats-new/comm-meet/2008/101608/E-1.pdf>

- [15] —, “Order No. 745, Demand Response Compensation in Organized Wholesale Energy Markets,” March 2011. [Online]. Available: <http://www.ferc.gov/EventCalendar/Files/20110315105757-RM10-17-000.pdf>
- [16] C. Chen, S. Kishore, and L. Snyder, “An Innovative RTP-based Residential Power Scheduling Scheme for Smart Grids,” in *Proc. of IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2011.
- [17] D. O’Neil, M. Levorato, A. Goldsmith, and U. Mitra, “Residential Demand Response Using Reinforcement Learning,” in *Proc. of IEEE International Conference on Smart Grid Communications*, 2010.
- [18] P. Du and N. Lu, “Appliance Commitment for Household Load Scheduling,” *IEEE Trans. on Smart Grid*, vol. 2, no. 2, pp. 411–419, 2011.
- [19] T. Hubert and G. Grijalva, “Realizing Smart Grid Benefits Requires Energy Optimization Algorithms at Residential Level,” in *Proc. of IEEE Conference on Innovative Smart Grid Technologies (ISGT)*, 2010.
- [20] D. Livengood, “The Energy Box: Comparing Locally Automated Control Strategies of Residential Electricity Consumption under Uncertainty,” *PhD Dissertation, Massachusetts Institute of Technology*, 2011.
- [21] R. Yin, P. Xu, M. Piette, and S. Kiliccote, “Study on Auto-DR and Pre-cooling of Commercial Buildings with Thermal Mass in California,” *Energy and Buildings*, vol. 42, no. 7, pp. 967–975, 2010.
- [22] Y. Zong, D. Kullmann, A. Thavlov, O. Gehrke, and H. Bindner, “Application of Model Predictive Control for Active Load Management in a Distributed Power System With High Wind Penetration,” *IEEE Trans. on Smart Grid*, vol. 3, no. 2, pp. 1055–1062, 2012.

- [23] F. Oldewurtel, A. Ulbig, A. Parisio, G. Andersson, and M. Morari, “Reducing Peak Electricity Demand in Building Climate Control using Real-Time Pricing and Model Predictive Control,” in *Proc. of IEEE Conference on Decision and Control*, 2010.
- [24] J. Contreras, R. Espinola, F. Nogales, and A. Conejo, “ARIMA Models to Predict Next-Day Electricity Prices,” *IEEE Trans. on Power Systems*, vol. 18, no. 3, pp. 1014–1020, 2003.
- [25] Z. Tan, J. Zhang, J. Wang, and J. Xu, “Day-Ahead Electricity Price Forecasting Using Wavelet Transform Combined with ARIMA and GARCH Models,” *Applied Energy*, vol. 87, no. 11, pp. 3606–3610, 2010.
- [26] C. Liu, J. Wang, A. Botterud, Y. Zhou, and A. Vyas, “Assessment of Impacts of PHEV Charging Patterns on Wind-Thermal Scheduling by Stochastic Unit Commitment,” *IEEE Trans. on Smart Grid*, vol. 3, no. 2, pp. 675–683, 2012.
- [27] Wikipedia, “Electric Vehicle Battery – Battery type: Lead-acid,” 2012. [Online]. Available: [http://en.wikipedia.org/wiki/Electric\\_vehicle\\_battery](http://en.wikipedia.org/wiki/Electric_vehicle_battery)
- [28] Battery University - Cadex Electronics Inc., “How to Prolong Lithium-based Batteries,” 2012. [Online]. Available: [http://batteryuniversity.com/learn/article/how\\_to\\_prolong\\_lithium\\_based\\_batteries](http://batteryuniversity.com/learn/article/how_to_prolong_lithium_based_batteries)
- [29] R. Hendron and C. Engebrecht, “Building America Research Benchmark Definition,” National Renewable Energy Laboratory (NREL), Tech. Rep., Jan 2010.
- [30] The Garlinghouse Company, “Multi-Family Plan 94481.” [Online]. Available: [http://www.familyhomeplans.com/plan\\_details.cfm?PlanNumber=94481](http://www.familyhomeplans.com/plan_details.cfm?PlanNumber=94481)

- [31] G. Argenbroe, J. Brown, Y. Heo, S. Kim, Z. Li, S. McManus, and F. Zhao, “Lessons from an Advanced Building Simulation Course,” in *Third National Conference of IBPSA-USA*, Berkeley, CA, 2008.
- [32] U.S. Department of Energy, “EnergyPlus Energy Simulation Software – Weather Data,” 2012. [Online]. Available: [http://apps1.eere.energy.gov/buildings/energyplus/cfm/weather\\_data.cfm](http://apps1.eere.energy.gov/buildings/energyplus/cfm/weather_data.cfm)
- [33] International Organization for Standardization (ISO), “Ergonomics of the Thermal Environment – Analytical Determination and Interpretation of Thermal Comfort using Calculation of the PMV and PPD Indices and Local Thermal Comfort Criteria,” *ISO 7730 Standard, Third Edition*, Nov. 2005.
- [34] ASHRAE, “2009 ASHRAE Handbook – Fundamentals,” American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Tech. Rep., 2009.
- [35] PJM Interconnection, “Day-ahead LMP Data,” 2011. [Online]. Available: <http://pjm.com/markets-and-operations/energy/day-ahead/lmpda.aspx>
- [36] New York ISO, “New York ISO Day-ahead Locational-Based Marginal Price (LBMP)-Zonal, Archive,” 2011. [Online]. Available: <http://mis.nyiso.com/public/P-2Alist.htm>
- [37] —, “New York ISO Real Time Market Locational-Based Marginal Price (LBMP)-Zonal, Archive,” 2011. [Online]. Available: <http://mis.nyiso.com/public/P-24Alist.htm>
- [38] Nissan, “Nissan Leaf – Features and Specifications,” 2012. [Online]. Available: <http://www.nissanusa.com/leaf-electric-car/index>
- [39] IBM, “IBM ILOG CPLEX Optimizer,” 2012. [Online]. Available: [www-01.ibm.com/software/integration/optimization/cplex-optimizer](http://www-01.ibm.com/software/integration/optimization/cplex-optimizer)

- [40] Solar Direct, “U.S. Small Wind Tax Credit.” [Online]. Available: [http://www.solardirect.com/rebates/wind\\_skystream.html](http://www.solardirect.com/rebates/wind_skystream.html)
- [41] —, “Solar Installation and Service Solutions.” [Online]. Available: <http://www.solardirect.com/services/installations.html>
- [42] A. Mohsenian-Rad, V. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, “Autonomous Demand Side Management Based on Game-Theoretic Energy Consumption Scheduling for the Future Smart Grid,” *IEEE Trans. on Smart Grid*, vol. 1, no. 3, pp. 320–331, 2010.
- [43] D. L. Ha, M. H. Le, and S. Ploix, “An Approach for Home Load Energy Management Problem in Uncertain Context,” *Proc. of IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 336–339, 2008.
- [44] P. Samadi, H. Mohsenian-Rad, R. Schober, V. Wong, and J. Jatskevich, “Optimal Real-time Pricing Algorithm Based on Utility Maximization for Smart Grid,” in *Proc. of IEEE Conf. on Smart Grid Communications*, 2010.
- [45] H. Kellerer, U. Pferschy, and D. Pisinger, *Knapsack Problems*. Berlin: Springer, 2004.
- [46] G. Xiong, C. Chen, S. Kishore, and A. Yener, “Smart (In-Home) Power Scheduling for Demand Response on the Smart Grid,” in *Proc. of IEEE PES Innovative Smart Grid Technologies (ISGT)*, Anaheim, CA, 2011.
- [47] NREL Solar Radiation Research Laboratory, “Custom Monthly Wind Speed Data.” [Online]. Available: [http://www.nrel.gov/midc/srrl\\_bms](http://www.nrel.gov/midc/srrl_bms)
- [48] Raum Energy, “3.5 kW Wind Turbine System Specification Sheet.” [Online]. Available: <http://www.raumenergy.com/turbine3kw.html>

- [49] G. Papaefthymiou and B. Klockl, “MCMC for Wind Power Simulation,” *IEEE Transactions on Energy Conversion*, vol. 23, no. 1, pp. 234–240, 2008.
- [50] M. Albadi and E. El-Saadany, “Demand Response in Electricity Market: An Overview,” in *Proc. of IEEE Power Engineering Society General Meeting*, 2007.
- [51] A. Mohsenian-Rad and A. Leon-Garcia, “Optimal Residential Load Control With Price Prediction in Real-Time Electricity Pricing Environments,” *IEEE Trans. on Smart Grid*, vol. 1, no. 2, pp. 120–133, 2010.
- [52] A. Mohsenian-Rad, V. W. Wong, J. Jatskevich, and R. Schober, “Optimal and Autonomous Incentive-based Energy Consumption Scheduling Algorithm for Smart Grid,” in *Proc. of IEEE Innovative Smart Grid Technologies*, 2010.
- [53] L. He and L. Snyder, “A Bilevel Model for Retail Electricity Pricing with Flexible Loads,” *to be submitted to Journal of Energy Engineering*, 2013.
- [54] M. Osborne, *An Introduction to Game Theory*. Oxford University Press, 2003.
- [55] L. Chen, N. Li, S. Low, and J. Doyle, “Two Markets Models for Demand Response in Power Networks,” in *Proc. of IEEE Conf. on Smart Grid Communications*, 2010.
- [56] N. Li, L. Chen, and S. Low, “Optimal Demand Response Based on Utility Maximization in power networks,” in *Proc. of IEEE PES General Meeting*, 2011.
- [57] M. Alizadeh, A. Scaglione, and R. Thomas, “From Packet to Power Switching: Digital Direct Load Scheduling,” *IEEE Journal on Selected Areas in Commu-*

- nications: Smart Grid Communications Series*, vol. 30, no. 6, pp. 1027–1036, 2012.
- [58] V. Ranade and J. Beal, “Distributed Control for Small Customer Energy Demand Management,” in *Proc. IEEE International Conference on Self-Adaptive and Self-Organizing Systems*, 2010, pp. 11–20.
- [59] R. Saber and M. Murray, “Consensus Protocols for Networks of Dynamic Agents,” in *Proc. 2003 American Control Conference*, 2003, pp. 951–956.
- [60] R. Saber and R. Murray, “Consensus Problems in Networks of Agents with Switching Topology and Time-Delays,” *IEEE Trans. Automation Control*, no. 9, pp. 1520–1533, 2004.
- [61] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1985.
- [62] L. Xiao and S. Boyd, “Fast Linear Iterations for Distributed Averaging,” *Systems & Control Letters*, pp. 65–78, 2004.
- [63] R. Saber, J. Fax, and R. Murray, “Consensus and Cooperation in Networked Multi-Agent Systems,” *Proceedings of the IEEE*, no. 1, pp. 215–233, 2007.
- [64] R. Saber, “Ultrafast Consensus in Small-World Networks,” in *Proc. 2005 American Control Conference*, 2005, pp. 2371–2378.
- [65] B. O’Hara and A. Petrick, *IEEE 802.11 Handbook - A Designer’s Companion*. IEEE Standards Association, 2005.
- [66] OkSolar.com, “Typical Power Consumption,” 2012. [Online]. Available: <http://www.oksolar.com/technical/consumption.html>

- [67] New York ISO, “New York ISO Real Time Actual Load, Archive,” 2011. [Online]. Available: <http://mis.nyiso.com/public/P-58Blist.htm>
- [68] S. Borenstein, J. Bushnell, and C. Knittel, “Market Power in Electricity Market: Beyond Concentration Measures,” *Energy Journal*, vol. 20, no. 4, 1999.
- [69] B. Bobbs, “Linear Complementarity Models of Nash-Cournot Competition in Bilateral and POOLCO Power Market,” *Utilities Policy*, vol. 5, no. 3/4, pp. 194–202, 1996.
- [70] J. Yao, I. Adler, and S. Oren, “Modeling and Computing Two-settlement Oligopolistic Equilibrium in a Congested Electricity Network,” *Operations Research*, vol. 56, no. 1, pp. 34–47, 2008.
- [71] S. Borenstein, J. Bushnell, E. Kahn, and S. Stoft, “Market Power in California Electricity Market,” *IEEE trans. on Power Systems*, vol. 16, no. 2, 2001.
- [72] S. Borenstein, J. Bushnell, and C. Knittel, “An Empirical Analysis of the potential for market power in a Deregulated California Electricity Market,” *Journal of Industrial Economics*, vol. 47, 1999.
- [73] J. Yao, S. Oren, and I. Adler, “Cournot Equilibria in Two-settlement Electricity Markets with System Contingencies,” *Int. J. Critical Infrastructure*, vol. 3, no. 1/2, pp. 142–160, 2007.
- [74] —, “Two-settlement Electricity Markets with Price Caps and Cournot Generation Firms,” *European Journal of Operations Research*, vol. 181, pp. 1279–1296, 2007.
- [75] —, “Computing Two Settlement Cournot Equilibria in Electricity Markets,” in *Proc. of the 37th Hawaii International Conference on System Sciences*, 2004.

- [76] P. Klemperer and M. Meyer, “Supply Function Equilibrium in Oligopoly under Uncertainty.” *Econometrica*, vol. 57, no. 6, pp. 1243–1277, 1989.
- [77] R. Green and D. Newbery, “Competition in the British Electricity Market,” *Journal of Political Economy*, vol. 100, no. 5, pp. 929–953, 1992.
- [78] R. Baldick and W. Hogan, “Capacity Constrained Supply Function Equilibrium Models of Electricity Markets: Stability , Non-decreasing constraints, and Function Space Iterations,” *POWER Working Paper PWP-089, University of California Energy Institute*, 2001.
- [79] C. Chen, S. Kishore, Z. Wang, M. Alizadeh, and A. Scaglione, “A Cournot Game Analysis on Market Effects of Queuing Energy Request as Demand Response,” in *Proc. of IEEE PES General Meeting*, 2012.
- [80] A. Khodaei, M. Shahidehpour, and S. Bahramirad, “SCUC with Hourly Demand Response Considering Intertemporal Load Characteristics,” *IEEE Trans. on Smart Grid*, vol. 2, no. 3, pp. 564–571, 2011.
- [81] M. Parvania and M. Fotuhi-Firuzabad, “Demand Response Scheduling by Stochastic SCUC,” *IEEE Trans. on Smart Grid*, vol. 1, no. 1, pp. 89–98, 2010.
- [82] E. Karangelos and F. Bouffard, “Towards Full Integration of Demand-Side Resources in Joint Forward Energy/Reserve Electricity Markets,” *IEEE Trans. on Power Systems*, vol. 27, no. 1, pp. 280–289, 2012.
- [83] W. Schill and C. Kemfert, “The Effect of Market Power on Electricity Storage Utilization: The Case of Pumped Hydro Storage in Germany,” *The Energy Journal*, no. 3, pp. 59–87, 2011.

- [84] R. Bessa, M. Matos, F. Soares, and J. Lopes, “Optimized Bidding of a EV Aggregation Agent in the Electricity Market,” *IEEE Trans. on Smart Grid*, no. 1, pp. 443–452, 2012.
- [85] D. Wu, D. Aliprantis, and L. Ying, “Load Scheduling and Dispatch for Aggregators of Plug-in Electric Vehicles,” *IEEE Trans. on Smart Grid*, no. 1, pp. 368–376, 2012.
- [86] Q. Zhu, J. Zhang, and P. Sauer, “A Game-Theoretic Framework for Control of Distributed Renewable-Based Energy Resources in Smart Grids,” in *Proc. of 2012 American Control Conference (ACC)*, Montreal, Canada, 2012.
- [87] W. Saad, Z. Han, H. Poor, and T. Basar, “Game Theoretic Methods for the Smart Grid,” *IEEE Signal Processing Magazine, Special issue on Signal Processing for the Smart Grid*, vol. 29, no. 5, pp. 86–105, 2012.
- [88] D. Veit, A. Weidlich, J. Yao, and S. Oren, “Simulating the Dynamics in Two-settlement Electricity Markets via an Agent-based Approach,” *International Journal of Management Science and Engineering Management*, vol. 1, no. 2, pp. 83–97, 2006.
- [89] J. Rosen, “Existence and Uniqueness of Equilibrium Points for Concave  $n$ -person Games,” *Econometrica*, pp. 347–351, 1965.
- [90] R. Zimmerman, C. Murillo-Sanchez, and R. Thomas, “MATPOWER: Steady-state Operations, Planning and Analysis Tools for Power Systems Research and Education,” *IEEE Trans. on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

- [91] M. Alizadeh, A. Scaglione, J. Davies, and K. Kurani, “A Scalable Stochastic Model for the Electricity Demand of Electric and Plug-in Hybrid Vehicles,” *submitted to IEEE Trans. on Smart Grid*, 2012.

# Vita

Chen Chen was born in 1983 in Xi'an, Shaanxi Province, P.R. China. He received the B.Eng. degree in Information Engineering from Xi'an Jiaotong University, Xi'an, China in 2006, and the M.S. degree in Communication and Information System from Xi'an Jiaotong University, Xi'an, China in 2009. Since 2009, he has been working towards the Ph.D. degree in the Department of Electrical and Computer Engineering, Lehigh University, Bethlehem, PA.

From May 2012 to October 2012, he was a research aide intern at Decision and Information Sciences Division, Argonne National Laboratory, IL. His research interests include communications and signal processing in smart grid, with focus on demand response for energy management.

Mr. Chen Chen received the Rossin College of Engineering and Applied Science (RCEAS) Fellowship from Lehigh University in 2009, and several scholarships from Xi'an Jiaotong University from 2003 to 2008.

## List of Publications

### Journal Papers

1. C. Chen, K.G. Nagananda, G. Xiong, S. Kishore, L.V. Snyder, “A Communication Based Appliance Scheduling Scheme for Consumer-Premise Energy Management Systems”, *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 56-65, 2013.
2. C. Chen, J. Wang, Y. Heo, S. Kishore, “MPC-based Appliance Scheduling for Residential Building Energy Management Controller”, submitted to *IEEE Transactions on Smart Grids*, 2012, under review.
3. C. Chen, J. Wang, S. Kishore, “A Distributed Direct Load Control Approach for Large-Scale Residential Demand Response”, 2013, in preparation.
4. C. Chen, P. Venkitasubramaniam, S. Kishore, L. He, L.V. Snyder, “Achievable Privacy in Aggregate Energy Management Systems”, 2013, in preparation for *Journal of Energy Engineering*.
5. C. Chen, M. Alizadeh, Z. Wang, S. Kishore, A. Scaglione, “Effects of Demand Response from Aggregated Deferrable Loads in the Electricity Market”, 2013, in preparation for *IEEE Transactions on Power System*.

### Conference Papers

1. C. Chen, K.G. Nagananda, G. Xiong, S. Kishore, L.V. Snyder, “Analysis of a Joint Access and Scheduling Scheme for Residential Energy Management Controller”, *IEEE Sensor Array and Multichannel Signal Processing Workshop*, Invited Paper, Hoboken, NJ, June 2012.

2. C. Chen, S. Kishore, Z. Wang, M. Alizadeh, A. Scaglione, “A Cournot Game Analysis on Markets Effects of Queuing Energy Request as Demand Response”, *IEEE PES Annual General Meeting*, San Diego, CA, July 2012.
3. M. Alizadeh, Z. Wang, A. Scaglione, C. Chen, S. Kishore “On the Market Effects of Queueing Energy Requests as an Alternative to Storing Electricity”, *IEEE PES Annual General Meeting*, San Diego, CA, July 2012.
4. C. Chen, S. Kishore, Z. Wang, M. Alizadeh, and A. Scaglione, “How will Demand Response Aggregators Affect Electricity Markets? - A Cournot Game Analysis”, *5th International Symposium on Communications, Control, and Signal Processing (ISCCSP)*, Invited Paper, Rome, Italy, May 2012.
5. M. Alizadeh, T.H. Chang, A. Scaglione, C. Chen, S. Kishore, “The Emergence of Deferrable Energy Requests and a Greener Future: What Stands in the Way?”, *5th International Symposium on Communications, Control, and Signal Processing (ISCCSP)*, Invited Paper, Rome, Italy, May 2012.
6. C. Chen, S. Kishore, L.V. Snyder, “An Innovative RTP-Based Residential Power Scheduling Scheme for Smart Grids”, *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011.
7. G. Xiong, C. Chen, S. Kishore, A. Yener, “Smart (In-Home) Power Scheduling for Demand Response on the Smart Grid”, *IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*, Anaheim, CA, Jan. 2011.
8. C. Chen, S. Kishore, “Making Cooperation Valuable: A Delay-Centric, Pricing-based User Cooperation Strategy”, in *Proceedings of 45th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, March 2011.

## Journal Papers in Chinese

1. C. Chen, T. Zhang, “Design and Application of Viterbi Decoder in DSP based on VCP”, *Microcomputer Information* , no. 4, pp. 202-203, 2009.
2. C. Chen, L. Sun, T. Zhang, “The Design of an Adaptive-OFDM System and its Implementation on DSP”, *Microelectronics and Computer*, vol. 26, no. 5, pp. 1-5, 2009.
3. L. Sun, C. Chen, W. Deng, R. Qiao, “Design and Implementation of Baseband Signal Processing Module for OFDM System Based on Software Defined Radio”, chapter in *The Prize Achievement Assembly of TI DSP Contest in 2008*, Publishing House of Electronic Industry, 2008.