1-28-2015

Machine Learning Aided Decision Making and Adaptive Stochastic Control in a Hierarchical Interactive Smart Grid

Ding Li

Ding Li

Candidate

Electrical and Computer Engineering

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. Sudharman K. Jayaweera , Chairperson

Dr. Chaouki T. Abdallah

Dr. Edward D. Graham

Dr. Francesco Sorrentino

Machine Learning Aided Decision Making and Adaptive Stochastic Control in a Hierarchical Interactive Smart Grid

by

Ding Li

B.E., Communication Engineering, Harbin Institute of Technology, 2008

M.S., Electrical Engineering, University of New Mexico, 2012

DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Engineering

The University of New Mexico

Albuquerque, New Mexico

December, 2014

Dedication

To my mother, Xin, for her encouragement and support.

"The nice part about being a pessimist is that you are constantly being either proven right or pleasantly surprised."

– George F. Will

Acknowledgments

First of all, I would like to thank my adviser, Prof. Sudharman K. Jayaweera, for his support and guidance over the past years at UNM. Prof. Jayaweera has taught me not only how to do research, but how to think independently as well. Working with Prof. Jayaweera has been a great learning experience for me. It is not easy to meet his standards for Ph.D. students. I have made it and I am proud of it. I would like also to thank my committee members, Professor Chaouki T. Abdallah, Professor Edward D. Graham and Professor Francesco Sorrentino, for their help and support. I really appreciate their helpful advices on my Ph.D. dissertation.

I would like to thank all my lab colleagues for their collaboration on research. It has been a great experience working with these talented researchers in the lab, and it is equally enjoyable playing basketball on the court with these not-that-talented players.

In the last, I would like to say "Thank you!" to my mom for her constant support. She raises me up to more than I can be. Moreover, in memory of my father, who passed away 16 years ago, I would like to thank him because his words have always been the source of strength and courage for me to keep moving forward. Life makes me a man like him. I did not let him down and I never will.

Machine Learning Aided Decision Making and Adaptive Stochastic Control in a Hierarchical Interactive Smart Grid

by

Ding Li

B.E., Communication Engineering, Harbin Institute of Technology, 2008

M.S., Electrical Engineering, University of New Mexico, 2012

Ph.D., Engineering, University of New Mexico, 2014

Abstract

In this dissertation, a hierarchical interactive architecture for the future smart grid is proposed. This hierarchical architecture consists of different layers ranging from the households, microgrid controller level, feeder level and substation level. The proposed smart grid architecture is scalable while allowing for sufficient resource pooling, because in each layer the power generation and consumption sides interact in a similar manner. Therefore, we develop an abstract Grid model with distributed energy resources (DER) and storage facilities. A comprehensive real time interactive scheme is proposed for the abstract Grid model, which addresses several important topics: (1) load prediction and uncertainty modeling, (2) demand response (DR), (3) stochastic tracking control of the conventional generation in the presence of DER's (both renewable energy and plug-in hybrid electric vehicle (PHEV)) and (4) machine learning aided decision making for smart-homes.

In the first part, a series of linear prediction models are presented for the load prediction purposes, including standard autoregressive (AR) process and time varying autoregressive (TVAR) process, according to different assumptions on the stationarity of customer load profile: piecewise stationarity, local stationarity and cyclostationarity. Two important issues in AR/TVAR models are addressed: determining the order of AR/TVAR models and calculating the AR/TVAR coefficients. The partial autocorrelation function (PACF) is analyzed to determine the model order and the minimum mean squared error (MMSE) estimator is adopted to derive the AR/TVAR coefficients, which leads to the Yule-Walker type of equations. In the second part, a DR scheduling scheme based on the Utility cost minimization with different customer clustering sizes. A convex optimization problem is formulated and the optimal demand response profile is in the form of a two-dimensional *water-filling solution* either with flat water levels or different water levels for different customers. A trade-off strategy which attempts to balance the competing objectives (centralized and distributed) is also provided based on the Price of Anarchy (PoA) analysis. In the third part, two stochastic tracking schemes are proposed to balance the power generation and consumption: (1) reference dynamics-based tracking and (2) reference statistics-based tracking. The proposed optimal tracking control schemes are further generalized by considering the realistic scenario with asynchronous net load demand signals from different customers. Based on the separation principle in reference prediction and tracking design, we propose both centralized and distributed reference prediction schemes based on Kalman filtering technique. In the forth part, with the hierarchical architecture well developed, the smart-home decision making problem is addressed by combining solutions to two sub-problems: (1) a hidden mode Markov decision process (HM-MDP) model based centralized sequential decision making at the microgrid controller to maximize an accumulated reward of the whole microgrid and (2) distributed auctioning game design among all smart-homes within the microgrid to coordinate their interactions based on the optimal energy decisions obtained

in the centralized sub-problem.

Contents

List of Figures

Publications

In the following, we present a list of journal and conference publications that have resulted from the work in this dissertation.

Book Chapter

1. D. Li and S. K. Jayaweera, "Machine-learning Aided Optimal Customer Decisions in Interactive Smart Grids", in *Smart Grids: Technologies, Applications and Management Systems*. Nova Science, 2014. Accepted for publication.

Journal Publications

1. D. Li and S. K. Jayaweera, "Distributed Smart-home Decision-making in a Hierarchical Interactive Smart Grid Architecture", *IEEE Transactions on Parallel and Distributed Systems*. Jan. 2014. Accepted for publication.

2. D. Li and S. K. Jayaweera, "Machine-learning Aided Optimal Customer Decisions for an Interactive Smart Grid", *IEEE Systems Journal*. Jun. 2014. Accepted for publication.

3. D. Li and S. K. Jayaweera, "Optimal Stochastic Tracking for Primary Frequency Control in an Interactive Smart Grid Infrastructure", *IEEE Systems Journal*, Jul. 2014. Accepted for publication.

4. D. Li and S. K. Jayaweera, "Uncertainty Modeling and Price-based Demand Response Scheduling in Smart Grid", *IEEE Systems Journal*, Apr. 2014, second round review.

Conference Publications

1. D. Li and S. K. Jayaweera, "Reinforcement Learning aided Smart-home Decision-making in an Interactive Smart Grid", 2014 *IEEE Green Energy and Systems Conference*, Mar. 2014. Accepted for publication.

2. D. Li and S. K. Jayaweera, "Uncertainty Modeling and Prediction for Customer Load Demand in Smart Grid", *2013 IEEE EnergyTech Conference*, May 2013, Cleveland, Ohio, USA.

3. S. K. Jayaweera and D. Li, "Reference Prediction in Optimal Control of Smart-grid with Asynchronous Distributed Renewables", *the 4th International Conference on Intelligence and Advanced Systems (ICIAS2012)*, Kuala Lumpur, Malaysia, Jun. 2012.

4. D. Li, S. K. Jayaweera and C. T. Abdallah, "Uncertainty Modeling and Stochastic Control Design for Smart Grid with Distributed Renewables", *IEEE Green Technologies Conference*, Tulsa, Oklahoma, USA, Apr. 2012.

5. D. Li, S. K. Jayaweera and A. Naser, "Auctioning game based Demand Response scheduling in smart grid," *2011 IEEE Online Conference on Green Communications (GreenCom11)*, Sept. 2011.

6. D. Li, S. K. Jayaweera, O. Lavrova and R. Jordan "Load Management for Price-based Demand Response Scheduling - a Block Scheduling Model," *2011 International Conference on Renewable Energies and Power Quality (ICREPQ'11)*, Las Palmas de Gran Canaria, Spain, Apr. 2011.

Technical Report

1. S. K. Jayaweera and D. Li, "Stochastic Control for Smart Grid with Integrated Renewable Distributed Generators," Technical Report, University of New Mexico, Jan. 2012.

Glossary

ADP	Approximate dynamic programming
DR	Demand response
DER	Distributed energy resource
IP	Incremental pruning
ML	Machine learning
MMSE	Minimum mean square error
MLE	Maximum likelihood estimator
RDG	Renewable distributed generation
PHEV	Plug-in hybrid electrical vehicle
PWLC	Piecewise linear and convex
TVAR	Time varying autoregressive
PACF	Partial autocorrelation function
HM-MDP	Hidden mode Markov decision process
POMDP	Partially observable Markov decision process

Chapter 1

Introduction

The future smart grid will be intelligent, efficient, resilient and green [1], enhancing every facet of the electric system, including generation, transmission, distribution and consumption that will transform the current grid to one that functions more cooperatively, responsively and organically. In traditional electric grid planning, inefficiencies are abound on both the generation and demand side, i.e. the Utility and customers interact on a slow time scale and ineffectively. Such interaction is mainly due to the insufficient information exchange between the generating and the consuming sides. In addition, highly time-varying demand/consumption profiles mean that matching power demand is a difficult proposition. For example, the widely adopted fixed pricing scheme makes the customers indifferent about scheduling their load demand within a day. Thus, similar consuming patterns among customers make huge peaks in the overall load demand profile. The need for the power generator to meet peak demand (as opposed to average demand) to prevent blackouts in the current paradigm inherently creates gross inefficiencies and is extremely costly for the Utility companies. For example, the U.S. national load factor is about 55%, and 10% of generation and 25% of distribution facilities are used less than 400 hours per year, i.e., 5% of the time [2].

Proliferation of distributed energy resources (DER), in particular renewable distributed generation, provides great promise in significantly improving the efficiency of electricity distribution. However, as DER's proliferate to a significant fraction of the overall electric energy on the distribution network, without proper procedures integration may lead to highly imbalanced transient behaviors which may overwhelm current infrastructure not to mention outages and brown-outs. *In a future smart grid, a customer with renewable generation capability (such as PV panels and wind turbines) may use predictive strategies to optimize its energy demand requests over time and determine when to use, sell or store its own renewable generation, flexibly interacting with the electric-grid and other customers, as opposed to being a passive energy consumer as today.* The information shared among distributed nodes (customers) endowed with generation, storage and consumption attributes can result in a distributed decision and control framework that will lead to both overall energy and cost efficiencies. Realizing the full potential promised by smart grid concept, however, requires systematic design principles, a comprehensive protocol framework for interaction among distributed entities that make up the grid and robust and computationally efficient control and optimization algorithms. The importance of this real time interaction framework is expected to become even more significant as high penetration of renewable generations and PHEVs appear in generation side and consumption side separately. This is because the distributed nature of power demands, as well as the intermittence of renewable generation, make both load and generation profiles fluctuating over time and difficult to be matched with each other.

As renewable generation, which is mostly based on solar, wind and tidal resources, grows at a rapid pace, renewable distributed generation (RDG) becomes a necessary and desirable component of a cleaner energy future. The benefits of integration of RDG's into the grid do not stop simply as another power source. Having RDG's at customer customers enable them to be energy-efficient while also achieving cost-savings. Indeed, this is where the smart grid really turns out to be smart:

The customers may adopt machine-learning aided predictive strategies to optimize their energy demand requests to the Utility and determine when to use, sell or store its own renewable generation. The Utility, in turn, may benefit by scheduling its demand-response to operate cost-effectively while ensuring better electric quality to the customers. While RDG's provide an opportunity to aid in balancing a highly variable load, their inherently intermittent generation profile could potentially become a source of instability. There are many technical challenges to increased penetration of RDG, such as voltage rise effects, power quality and power grid protection when they are to be integrated into the traditional power grid (to form a smart grid). The most challenging aspect of integrating renewable distributed generators (renewables) is dealing with their inherent intermittent generation profile. Historically, in the equation of supply and demand, operators have primarily had to deal with the demand variable. With more integrated renewable distributed generators coming online, however, operators need more efficient and effective control schemes to balance variables on both sides of the equation [3].

Although a comprehensive formulation and an analysis is not yet available, still there have been some attempts to understand, model and analyze these effects [4, 5]. For example, a multi-stage frequency control framework is presented in [6–8]. However, it does not address the issue of consumption planning on the customer side. The uncertainty in supply due to integrated renewable DER's and the challenges they impose on the existing distribution infrastructure and the system operator have been discussed in [9]. The distribution-level smart grid features such as interconnection of distributed generation and active distribution management, automated meter reading (AMR) systems in network management and power quality monitoring were discussed in [10]. In [11], the implementation of vehicle-to-grid (V2G) power issues, strategies and business models for doing so, for purposes of both stabilizing the grid and supporting large-scale renewable energy were discussed. Various control-theoretic and system-level problem formulations of smart grid architectures

have been discussed in [12] and [13]. In [12], for example, the authors showed that significant improvements can be made to the operations of a smart grid by providing information about the likely behavior of renewable energy through both online short-term forecasting and longer-term assessments. In [13], a distributed control method was proposed for converter-interfaced renewable generation units with active filtering capability.

To fully harvest the fruits of RDG integration, the real time interaction between the customers and the Utility needs to be robust against possible communications, sensing and actuation delays and errors. Ultimately, of course, the Utility has to make sure that the power-grid is stable under this real time, distributed and networked interaction among RDG's, customer appliances and the conventional plant maintained by the Utility. In this dissertation, a comprehensive real-time interactive framework is developed for the Utility and customers in a smart grid while ensuring grid-stability and Quality-of-Service (QoS), as shown in Fig. 5.1.

This hierarchical architecture for the smart grid is scalable while allowing for sufficient resource pooling [14]. The scalability of the grid requires being able to easily integrate additional customers into the grid without affecting the established operational conditions of the grid. Ideally this might be achievable if each individual household is managed separately, but, of course, this would preclude any resource pooling, which is one of the most important strategies to energy efficiency in the grid. A tradeoff to this can be achieved by using the notion of microgrids with DER's. Each microgrid is a collection of households with certain self-containing capabilities, which are geographically adjacent and coordinated by a microgrid controller, as shown in the red box in Fig. 5.1. However, we can also think of each approximately self-contained microgrid as a broader customer unit coordinated by a feeder-level controller as shown in the blue box in Fig. 5.1. Similarly, we can scale up to the substation level and above and investigate an entire hierarchical smart grid architecture,

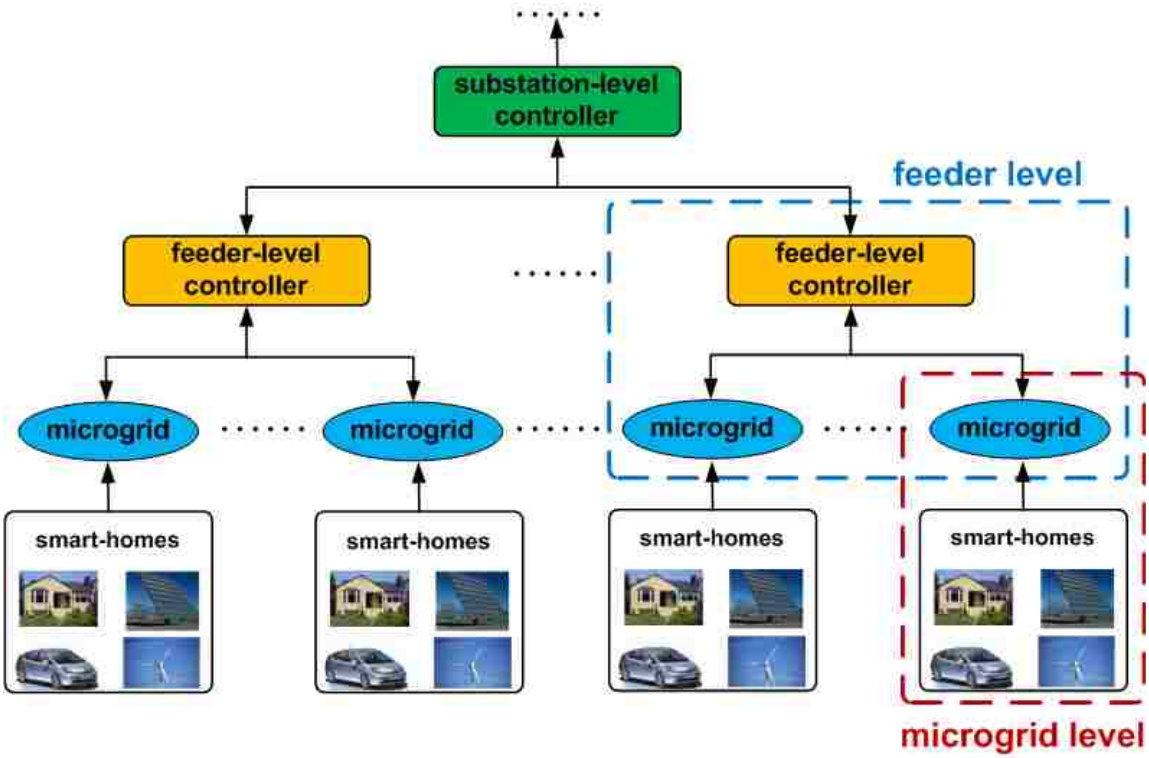

Figure 1.1: Hierarchical smart grid architecture that is scalable while allowing for sufficient resource pooling.

as shown in Fig. 5.1.

As we scale up to construct the entire grid, at each level, all branches with the same structure of one controller and multiple customer units are all approximately self-contained and are coordinated by the controller at a higher level. For example, at the microgrid level in Fig. 5.1, all microgrid branches identical to the red box are approximately self-contained. When the power-load mismatch is too big to be mitigated within a single microgrid, electric power flow will be routed among different microgrids under the coordination of a feeder-level controller. Similarly, at the feeder-level in Fig. 5.1, all branches identical to the blue box are approximately self-contained. Power flow among feeder-level branches are to be coordinated by the

substation-level controller. Hence, with this hierarchical architecture interpretation, any decision-making framework designed for a controller and the individual units below it is applicable to each of the levels in this hierarchical smart grid. Thus, in the following, we may focus on an abstract model made of a single (micro-grid) controller and a collection of multiple (smart-home) customers managed by it.

This hierarchical architecture for the Utility-customer interaction is made of the following sub-components of *customer load prediction*, *demand response*, *active power-load balancing* and *smart-home decision making*.

1.1 Uncertainty Modeling and Prediction for Customer Load Demand in Smart Grid

Precise prediction and modeling of the uncertainties has always been an important and challenging issue in power generation planning and load matching in electrical power grid [2]. In the upgrade from a traditional power grid to a "smarter" grid, which enables more efficiency and flexibility in grid operation, integrated renewable distributed generation (RDG) and plug-in hybrid electric vehicle (PHEV) are not only important features and driving forces, but also sources of uncertainties and instabilities. Hence, this challenging issue is expected to become more and more significant as the ever-increasing penetrations of RDG and PHEV appear on power generation and consumption sides separately. In future smart grid, the "smart home" is not only the households with "smart devices" such as advanced metering infrastructure (AMI) [15], but also distributed unit with local generation (on-site RDG's) and storage facilities (PHEV). Customers are capable of making optimal sequential decisions over a certain period of time, maximizing accumulated benefit based on the forecasted price information [16]. Note that the optimal sequential decisions are

made by the smart homes based on predictions of local generation and consumption information. Hence, appropriate models that capture the randomness and time transient features of both uncertainties are essential for making precise predictions.

The benefits of integration of RDG's into the grid do not stop simply as another power source. Having RDG's at customer premises enable them to be energy-efficient while also achieving cost-savings. A lot of studies have been reported in the literature proposing different stochastic models for renewable generation and related natural phenomena. For example, wind speed distributions are often characterized by Weibull or Rayleigh distributions [17, 18]. Historical hourly data for the wind farm site collected over a significant time are normally required to obtain the shaping parameters. In [18], the wind speed probability distributions obtained for the three diverse geographic locations in Canada, are close to normal distributions. The solar irradiation forecasting precision varies depending mainly on the quality of data in reference to the different dynamics of solar irradiation behavior. Beta distribution validated by different researches as a simple and sufficiently flexible two-parameter distribution, fits well the empirical data in many situations [19].

Different stochastic approaches have been reported in the literature for customer load modeling and two approaches are widely adopted. The first approach is component-based load modeling approach, which reconstructs the expected daily electrical loads of a household based on appliance sets, occupancy patterns, and statistical data. For example, in [20], the authors constructed such electric load profiles from individual appliance profiles. By considering "availability" and "proclivity" functions, they predict whether someone is available (at home and awake) and their tendency to use an appliance at any given time. These functions were applied to predict individual appliance events, which were then aggregated into a load profile. The second approach is termed the measurement-based load modeling. In [21], the authors used this approach to create electrical profiles to examine demand side man-

agement strategies for Finland. In [22], a methodology of measurement-based load modeling for transient stability analysis was proposed and Genetic Algorithms (GA) was used to estimate load model parameters. However, both approaches fail to give enough emphasis on the power consumption transition property over time.

In this dissertation, two types of approaches are proposed to model the uncertainty in customer load demand. The first approach is based on a first order non-stationary Markov chain. A maximum likelihood estimator (MLE) is derived to estimate the time variant transition matrix of the Markov chain. The second approach is based on time series analysis techniques. We present linear prediction models such as standard autoregressive (AR) process and time varying autoregressive process (TVAR), according to different assumptions on the stationarity of the customer load data: piecewise stationarity, local stationarity and cyclo-stationarity. Two important issues in AR/TVAR models are investigated: AR/TVAR coefficient estimation and determining the order of AR/TVAR models. The minimum mean squared error (MMSE) estimator is adopted to derive the AR/TVAR coefficients, which leads to the Yule-Walker type of equations. For the TVAR model, by doing basis function expansion based coefficient parametrization, we replace the scalar process with a vector one and turn the original non-stationary problem into a time-invariant problem. All the proposed models are tested against the same set of real measured customer load demand data. Prediction performances of different models are analyzed and compared, advantages and disadvantages are discussed. Both the non-stationary Markov chain and the linear prediction technique address the time transition property of the load demands.

1.2 Price-based Demand Response Scheme Design in Smart Grid

In most current electricity markets, fixed pricing schemes with constant rates are being widely used. Customers face retail electricity prices that are flat over months or even years [23]. A problem with fixed pricing schemes is the disconnection between short-term marginal electricity production costs and retail rates paid by customers, which leads to inefficient overall resource usage. Due to lack of information on generation costs, electricity consumption behavior of customers may not adjust to supply-side conditions. Thus fixed constant pricing results in suboptimal customer behavior as well as higher electricity costs than they would otherwise be in an optimally efficient system [24].

There is a growing consensus that Demand Response (DR) can play an important role in market design [25]. Lack of DR has been shown to be a major contributing factor for energy-market meltdowns [26]. In [23], for example, 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."* DR not only reduces the capacity investments in peak generation units to serve occasional heightened demand, but also provides short-term reliability benefits as it can offer load relief to resolve system and local capacity constraints. There are two basic demand response options: Price-based demand response and incentive-based demand response. Price-based demand response includes real-time pricing (RTP), critical-peak pricing (CPP), and time-of-use (TOU) rates. Customers can respond to the price structure with changes in energy use, reducing their electricity bills if they adjust the timing of their electricity usage to take advantage of lower-priced periods and avoid consuming when prices

are higher [23]. Incentive-based demand response schemes pay participants to reduce their loads at times requested by the program sponsor, triggered either by a grid reliability problem or high electricity prices. DR programs typically specify a method for establishing customers baseline energy consumption level below which demand reductions are not allowed. In power systems, the energy requests that customers send to utility consist of two parts: nonflexible load request and flexible load request [27]. The nonflexible part is the minimum amount of energy that utility needs to provide at a specific time. The flexible part can be reallocated over time according to a certain load management strategy. For any load management strategy there are two common primary goals: peak load shaving and load profile flattening. Under real-time pricing, the electricity price is determined by real time load information.

In this dissertation, a block scheduling model is presented for price-based demand response scheduling. In this model, the size of the time block is set to be small enough so that all load shifting within the time block can be considered as cost free and acceptable to customers. The solution to this block processing problem can then be the basis for implementations of arbitrarily long scheduling periods. Two types of real-time pricing schemes, linear pricing and threshold pricing, are discussed in this paper. We consider optimal demand-response when customers cooperate as a group as well as when each customer is only interested in minimizing its own cost. Naturally these two scenarios, as shown to lead to centralized and distributed optimizations.

1.3 Optimal Stochastic Tracking for Primary Frequency Control in Smart Grid

In any electric system, the stability of the electrical grid is guaranteed by balancing the power generation and consumption [28]. Generation units and even load in some

cases must be manipulated to conduct power balancing so the network user is not affected by load changes or generation and transmission outages. From the viewpoint of load matching, various demand response schemes have been proposed to affect customer load profiles [23, 29–31]. In [32], a three-step methodology was presented to manage the cooperation among technologies of distributed generation, distributed electricity storage and demand side load management. From the viewpoint of power generation control, since massive storage of alternating electricity is difficult, two separate equilibria should be kept on the grid for stabilizing purpose [33]: (1) The active power generated should at each moment equal the active power consumed. A deviation from this equilibrium results in a deviation from the standard frequency (60 or 50 Hz). Hence, keeping this equilibrium between active power consumption and generation means maintaining frequency. (2) The reactive power on the grid should be kept in equilibrium as well. Reactive power is an extra load for the grid, leaving less capacity for active power, resulting in a local voltage drop. Hence, keeping reactive power in equilibrium means maintaining voltage. Studies on frequency and voltage control have been reported in many previous work [28, 34–37]. In particular, a comprehensive survey on frequency and voltage control technical features can be found in [28]. In [34], the authors discuss the issue of excess steady-state voltage rise and the methods of limitation that can be applied with specific reference to wind generation. In [36], a strategy for the control of terminal voltage and frequency of a stand-alone self-excited induction generator-(SEIG) based wind generator, working with variable speed and load is proposed. In [37], the authors presented a micro hydro scheme with parallel operation of synchronous and induction generators in micro hydro scheme.

In most of the literature, frequency and voltage control schemes are usually designed separately because generally they are implemented by generator rotor speed governor and excitation control system respectively [8, 38–41]. In this paper, we focus on frequency control (active power control) issues. The frequency control

usually consists several layers [6–8], including primary control, secondary control, tertiary control and other possible balancing power reserve planning services. Control schemes of different levels have different objectives and operating time scales, as shown in Fig. 4.1. The objective of primary frequency control (with a controlling period on the order of seconds) is to maintain a balance between generation and consumption within the synchronous area using turbine speed or turbine governors. However, primary frequency control stabilizes frequency but does not drive the system frequency back to the original set-point value after a disturbance. Secondary frequency control (with a controlling period on the order of minutes) is needed since when several generators are doing generation sharing, secondary frequency control distributes the power imbalance among selected units [6]. The secondary frequency control can also drive the system frequency back to the original desired value. Tertiary frequency control (with a controlling period in the order of minutes to hours) is a manual change in the dispatching in order to restore the secondary reserve and provide a more permanent solution if the imbalance between consumed power and scheduled power persists. There are several important research issues associated with both secondary and tertiary frequency control, such as *spinning reserve*, *unit commitment* and *economic dispatch*. Spinning reserve [42] is the unused capacity provided by devices that are synchronized to the network and can be quickly activated on decision of the system operator. Unit commitment and economic dispatch [43] is to find the optimal dispatch of available generation resources to meet the electrical load and spinning reserves. Other layers in the frequency control framework include stand-by supplies and contractual load shedding which have longer control periods (hours). Unit commitment and economic dispatch are important topics in power grid generation planning by themselves and will not be discussed in detail here.

Currently the most widely adopted primary frequency control scheme is the proportional-integral-derivative controller (PID controller) [6]. This is because PID controller shows relatively good control performance when the dynamics of the plant

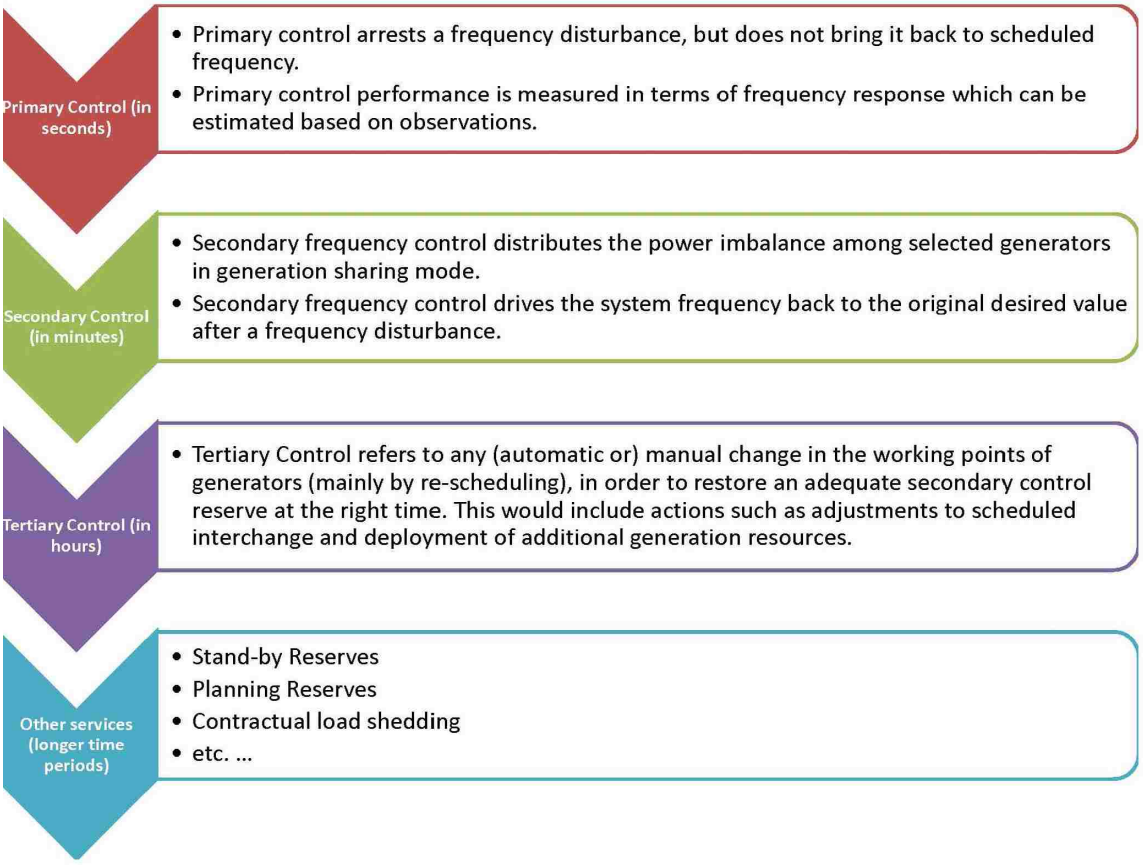

Figure 1.2: Frequency control consists primary control, secondary control, tertiary control and other planning reserve services.

is unknown or too complicated to analyze. By tuning the three parameters in the PID controller algorithm, the controller can conveniently provide control action designed for specific process requirements. However, PID controller does not guarantee optimality in control and system stability. In this paper, we focus on the primary frequency control design and propose an optimal stochastic tracking scheme for synchronous generator active power generation control, assuming the dynamics of individual synchronous generator. In this tracking scheme we minimize the difference between the active power generation output and the reference signal which incorporate the randomness of both load demands and renewable generations. Fur-

ther analysis on the tracking performance are presented considering synchronous and asynchronous customer load signals in the reference.

In this dissertation, an optimal stochastic tracking scheme is proposed in an interactive smart grid infrastructure consisting of three parts: (1) The Utility reshapes the customer load profiles by scheduling a demand response (DR) for the requested customer loads. (2) Individual smart-home makes optimal sequential decision on power purchase. (3) Optimal stochastic control schemes for the primary frequency control (active power control) are designed, in the presence of uncertainties arising from customer loads and distributed renewable generations. With the first two parts addressed in our previous work, in this paper, we focus on the primary frequency control scheme design in the multi-layer control architecture to stabilize frequency and maintain the active power balance within the distributed areas. We propose two stochastic tracking schemes based on the state-space representation of a synchronous generator: (1) reference dynamics-based tracking and (2) reference statistics-based tracking. We further extend the proposed optimal controllers by considering the realistic scenario of asynchronous load signals from different customers. To compensate for different delays seen by different customer signals, a Kalman filter (KF) based prediction scheme is proposed to estimate the correct reference signal. We show that the centralized reference prediction can equivalently be implemented distributively.

1.4 Machine-learning Aided Optimal Customer Decisions with an Auctioning Game Design for Interactive Smart Grids

Proliferation of distributed energy resources (DER), in particular renewable distributed generation, provides great promise in significantly improving the efficiency

of electricity distribution. However, as DER's proliferate to a significant fraction of the overall electric energy on the distribution network, without proper procedures integration may lead to highly imbalanced transient behaviors which may overwhelm current infrastructure not to mention outages and brown-outs. *In a future smart grid, a customer with renewable generation capability (such as PV panels and wind turbines) may use predictive strategies to optimize its energy demand requests over time and determine when to use, sell or store its own renewable generation, flexibly interacting with the electric-grid and other customers, as opposed to being a passive energy consumer as today.* The information shared among distributed nodes (customers) endowed with generation, storage and consumption attributes can result in a distributed decision and control framework that will lead to both overall energy and cost efficiencies. Realizing the full potential promised by smart grid concept, however, requires systematic design principles, a comprehensive protocol framework for interaction among distributed entities that make up the grid and robust and computationally efficient control and optimization algorithms.

The topic of customer decision making consists of several important subtopics, including smart-home design [44] and [45], system integration of distributed energy resources (DER) [46], renewable generation modeling [18,19,47], load demand modeling [20] and [21], and plug-in hybrid electrical vehicle (PHEV) vehicle-to-grid (V2G) management and regulation [48] and [49]. There is a considerable amount of previous studies reported on these subtopics in literature. For example, a smart-home energy management system based on a ZigBee sensor network was proposed in [44]. In [45], the author motivated the use of power line LANs as a basic infrastructure for building integrated smart homes, proposing protocols capable of supporting power line communication networks at speeds comparable to wired LANs. These smart-home models are mostly from the perspective of information gathering and transmission (e.g., a ZigBee sensor network and a power line LAN). However, it is unclear how these smart-home models can be evolved to allow real-time decision making that

makes use of all collected information. In [46], the concept of virtual power plant (VPP) was developed to enhance the control of DER by the system operators and other market actors by providing an appropriate interface between these system components. However, an equally important issue on the customer side (rather than the system operator side), which is the distributed self-management of DER's with local objectives, remains unaddressed. Various stochastic models for different renewable generations have been proposed in previous literature. For example, wind speed distributions are often characterized by Weibull or Rayleigh distributions [17]. The wind speed probability distributions obtained in [18] for three diverse geographic locations in Canada have been shown to be close to normal distributions. Beta distribution has been validated by different research as a simple and sufficiently flexible two-parameter distribution to fit the empirical solar irradiation behavior data in many situations [19]. These stochastic models are important, but these papers failed to present further discussions on how these models can be incorporated in customer decisions. Similar issue arises with papers focusing on customer load modeling and prediction, for example in [20] and [21]. In [48], the impact of charging PHEVs on a distribution transformer under different charging scenarios were examined. In [49], the author established a series of well-defined electric vehicle loads that were subsequently used to analyze their electric energy usage and storage in the context of more electrified road transportation. The PHEV management strategies mentioned above are part of the customer decision making addressed in this paper. However, it is important to consider more general energy decisions, rather than only focusing on PHEV charging strategy, taking into account of other factors, such as the impact of intermittent renewable generations. The work presented by the literature mentioned above provide an important foundation for upgrading the conventional grid-customer models to smart customers in a modern smart grid. However, little of these existing studies has considered a comprehensive cycle of interactions between the Utility and the distributed entities (customers) taking into account aspects of customer-side

decision making, Utility-side demand response scheduling, renewable DER integration and power-load balances for grid-stability and the effects of information and communication technology (ICT) infrastructure on all these.

With the hierarchical architecture in place, the concept of *smart-home* is extended in two aspects: (1) from traditional households with smart-devices, such as Advanced Metering Infrastructure (AMI), to intelligent entities with instantaneous and distributive decision making capabilities. (2) from individual households to general customer units of possibly large scales. We focus on the problem of real-time scheduling in an abstract grid model consisting of one microgrid controller and multiple smart-home customer units. A scalable solution to the real-time scheduling problem is proposed by combining solutions to two sub-problems: (1) centralized sequential decision making at the microgrid controller to maximize an accumulated reward of the whole microgrid and (2) distributed auctioning game design among all smart-homes to coordinate their interactions based on the optimal energy decisions obtained in the first centralized sub-problem.

For the centralized decision making problem at the microgrid controller, we adopt a hidden mode Markov decision process (HM-MDP) model. This real-time decision making framework can effectively be integrated with demand response (DR) schemes, which is prediction based and therefore inevitably leads to real-time power-load mismatches. With the Baum-Welch algorithm adopted to learn the non-stationary dynamics of the environment, we propose a value iteration (VI) based exact solution algorithm for the HM-MDP problem. Different from the conventional value iteration, the concept of *parsimonious sets* is used to enable a finite representation of the optimal value function. Instead of iterating the value function in each time step, we iterate the representational parsimonious sets by using the incremental pruning (IP) algorithm. Though this exact algorithm leads to optimal policies giving maximum rewards, its complexity suffers from the *curse of dimensionality*. To obtain

a low-complexity, real-time algorithm that allows adaptively incorporating new observations as the environment changes, we resort to Q-learning based approximate dynamic programming (ADP). Q-learning offers more flexibility in practice because it does not require specific starting and ending points of the scheduling period.

For the decentralized decision making problem at smart-homes within the microgrid, a Vikrey auctioning game is designed to coordinate the actions of the individual smart-homes to achieve the optimal solution derived by the microgrid controller in the centralized decision stage under realistic gird interaction assumptions. It is worth pointing out that application of different auction schemes for smart grid problems have been reported in [50–52]. For example, auction mechanisms that can be used by the aggregators for procuring stochastic renewable generations are proposed in [50]. In [51] and [52], double auction is adopted for distributed energy resources (DERs) management and Plug-in hybrid electric vehicles (PHEVs), respectively. However, most of these are focused on the solution derivation of auctions and fail to address the connection between the centralized and distributed decision schemes, which is important for the hierarchical architecture of the modern smart grid. In this dissertation, we show that though truthful bidding is a weakly dominant strategy for all smart-homes in the auctioning game, collusive equilibria do exist and can jeopardize the effectiveness and efficiency of the trading opportunity allocation. Analysis on the structure of the Bayesian Nash equilibrium solution set shows that the Vickrey auctioning game can be made more robust against collusion by customers (anticipating distributed smart-homes) by introducing a positive reserve price. The corresponding auctioning game is then shown to converge to the unique incentive compatible truthful bidding Bayesian Nash equilibrium, without jeopardizing the auctioneer's (microgrid controller's) profit. The performance analysis of both the proposed centralized and distributed decision making schemes are presented. This two-step solution approach is shown to be scalable to more complicated smart grid architectures beyond the assumed abstract model.

1.5 Dissertation Contributions

The main contributions of this dissertation can be summarized as follows:

- In this dissertation, two types of approaches are presented to model the uncertainty in customer load demand. The first approach is based on a first order non-stationary Markov chain. A maximum likelihood estimator (MLE) is derived to estimate the time variant transition matrix of the Markov chain. The second approach is based on time series analysis techniques. We present linear prediction models such as standard autoregressive (AR) process and time varying autoregressive process (TVAR), according to different assumptions on the stationarity of the customer load data: piecewise stationarity, local stationarity and cyclo-stationarity. Two important issues in AR/TVAR models are investigated: AR/TVAR coefficient estimation and determining the order of AR/TVAR models. The minimum mean squared error (MMSE) estimator is adopted to derive the AR/TVAR coefficients, which leads to the Yule-Walker type of equations. For the TVAR model, by doing basis function expansion based coefficient parametrization, we replace the scalar process with a vector one and turn the original non-stationary problem into a time-invariant problem. All the proposed models are tested against the same set of real measured customer load demand data. Prediction performances of different models are analyzed and compared, advantages and disadvantages are discussed. Both the non-stationary Markov chain and the linear prediction technique address the time transition property of the load demands.

- In this dissertation, a block scheduling model is presented for price-based demand response scheduling. In this model, the size of the time block is set to be small enough so that all load shifting within the time block can be considered as cost free and acceptable to customers. The solution to this block

processing problem can then be the basis for implementations of arbitrarily long scheduling periods. Two types of real-time pricing schemes, linear pricing and threshold pricing, are discussed in this paper. We consider optimal demand-response when customers cooperate as a group as well as when each customer is only interested in minimizing its own cost. Naturally these two scenarios, as shown to lead to centralized and distributed optimizations.

- In this dissertation, an optimal stochastic tracking scheme is proposed in an interactive smart grid infrastructure consisting of three parts: (1) The Utility reshapes the customer load profiles by scheduling a demand response (DR) for the requested customer loads. (2) Individual smart-home makes optimal sequential decision on power purchase. (3) Optimal stochastic control schemes for the primary frequency control (active power control) are designed, in the presence of uncertainties arising from customer loads and distributed renewable generations. With the first two parts addressed in our previous work, in this paper, we focus on the primary frequency control scheme design in the multi-layer control architecture to stabilize frequency and maintain the active power balance within the distributed areas. We propose two stochastic tracking schemes based on the state-space representation of a synchronous generator: (1) reference dynamics-based tracking and (2) reference statistics-based tracking. We further extend the proposed optimal controllers by considering the realistic scenario of asynchronous load signals from different customers. To compensate for different delays seen by different customer signals, a Kalman filter (KF) based prediction scheme is proposed to estimate the correct reference signal. We show that the centralized reference prediction can equivalently be implemented distributively.

- In this dissertation, a comprehensive real-time interactive framework is developed for the Utility and customers in a smart grid while ensuring grid-stability

and Quality-of-Service (QoS). First, we propose a hierarchical architecture for the Utility-customer interaction consisting of sub-components of *customer load prediction*, *renewable generation integration*, *power-load balancing* and *demand response (DR)*. With the hierarchical architecture developed, the concept of *smart-home* is extended in two aspects: (1) from traditional households with smart-devices, such as Advanced Metering Infrastructure (AMI), to intelligent entities with instantaneous and distributive decision making capabilities. (2) from individual households to general customer units of possibly large scales. We focus on the problem of real-time scheduling in an abstract grid model consisting of one microgrid controller and multiple smart-home customer units. A scalable solution to the real-time scheduling problem is proposed by combining solutions to two sub-problems: (1) centralized sequential decision making at the microgrid controller to maximize an accumulated reward of the whole microgrid and (2) distributed auctioning game design among all smart-homes to coordinate their interactions based on the optimal energy decisions obtained in the first centralized sub-problem.

1.6 Structure of the Dissertation

The remainder of this dissertation is organized as follows: Chapter 2 proposes two types of approaches are presented to model the uncertainty in customer load demand: (1) the first order non-stationary Markov chain model and (2) time series analysis technique based approach. In Chapter 3, a block scheduling model is presented for price-based demand response scheduling with the two dimensional water filling results analyzed in detail. In Chapter 4, an optimal stochastic tracking scheme is proposed in an interactive smart grid infrastructure with both synchronous and asynchronous reference signals. In Chapter 5, we present the machine learning aided smart-home

decision making scheme with an auctioning game design. Finally, we conclude the dissertation in Chapter 6.

Chapter 2

Uncertainty Modeling and Prediction for Customer Load Demand

2.1 Introduction

Precise prediction and modeling of the uncertainties has always been an important and challenging issue in power generation planning and load matching in electrical power grid [2]. In the upgrade from a traditional power grid to a "smarter" grid, which enables more efficiency and flexibility in grid operation, integrated renewable distributed generation (RDG) and plug-in hybrid electric vehicle (PHEV) are not only important features and driving forces, but also sources of uncertainties and instabilities. Hence, this challenging issue is expected to become more and more significant as the ever-increasing penetrations of RDG and PHEV appear on power generation and consumption sides separately. In future smart grid, the "smart home" is not only the households with "smart devices" such as advanced metering infras-

tructure (AMI) [15], but also distributed unit with local generation (on-site RDG's) and storage facilities (PHEV). Customers are capable of making optimal sequential decisions over a certain period of time, maximizing accumulated benefit based on the forecasted price information [16]. Note that the optimal sequential decisions are made by the smart homes based on predictions of local generation and consumption information. Hence, appropriate models that capture the randomness and time transient features of both uncertainties are essential for making precise predictions.

The benefits of integration of RDG's into the grid do not stop simply as another power source. Having RDG's at customer premises enable them to be energy-efficient while also achieving cost-savings. A lot of studies have been reported in the literature proposing different stochastic models for renewable generation and related natural phenomena. For example, wind speed distributions are often characterized by Weibull or Rayleigh distributions [17, 18]. Historical hourly data for the wind farm site collected over a significant time are normally required to obtain the shaping parameters. In [18], the wind speed probability distributions obtained for the three diverse geographic locations in Canada, are close to normal distributions. The solar irradiation forecasting precision varies depending mainly on the quality of data in reference to the different dynamics of solar irradiation behavior. Beta distribution validated by different researches as a simple and sufficiently flexible two-parameter distribution, fits well the empirical data in many situations [19].

Different stochastic approaches have been reported in the literature for customer load modeling and two approaches are widely adopted. The first approach is component-based load modeling approach, which reconstructs the expected daily electrical loads of a household based on appliance sets, occupancy patterns, and statistical data. For example, in [20], the authors constructed such electric load profiles from individual appliance profiles. By considering "availability" and "proclivity" functions, they predict whether someone is available (at home and awake) and their

tendency to use an appliance at any given time. These functions were applied to predict individual appliance events, which were then aggregated into a load profile. The second approach is termed the measurement-based load modeling. In [21], the authors used this approach to create electrical profiles to examine demand side management strategies for Finland. In [22], a methodology of measurement-based load modeling for transient stability analysis was proposed and Genetic Algorithms (GA) was used to estimate load model parameters. However, both approaches fail to give enough emphasis on the power consumption transition property over time.

In this study, we propose two types of approaches to model the uncertainty in customer load demand. The first approach is based on a first order non-stationary Markov chain. A maximum likelihood estimator (MLE) is derived to estimate the time variant transition matrix of the Markov chain. The second approach is based on time series analysis techniques. We present linear prediction models such as standard autoregressive (AR) process and time varying autoregressive process (TVAR), according to different assumptions on the stationarity of the customer load data: piecewise stationarity, local stationarity and cyclo-stationarity. Two important issues in AR/TVAR models are investigated: AR/TVAR coefficient estimation and determining the order of AR/TVAR models. The minimum mean squared error (MMSE) estimator is adopted to derive the AR/TVAR coefficients, which leads to the Yule-Walker type of equations. For the TVAR model, by doing basis function expansion based coefficient parametrization, we replace the scalar process with a vector one and turn the original non-stationary problem into a time-invariant problem. All the proposed models are tested against the same set of real measured customer load demand data. Prediction performances of different models are analyzed and compared, advantages and disadvantages are discussed. Both the non-stationary Markov chain and the linear prediction technique address the time transition property of the load demands.

The rest of the chapter is organized as follows: In section 2.2, we present the non-stationary Markov chain model. Linear prediction models based on the standard autoregressive linear prediction models according to piecewise stationary and locally stationary assumptions on the customer load data are presented in section 2.3. In section 2.4, we discuss the TVAR model based on the cyclo-stationary assumption which is more general and realistic than previous assumptions. In section 2.5, we analyze and compare the prediction performances of all the approaches proposed. The conclusions of this chapter are given in section 2.6.

2.2 First-order non-stationary Markov Chain based Model

2.2.1 Maximum Likelihood Estimation of Transition Matrix

Determining the transition matrix is one of the key issues in Markov chain modeling. It becomes more challenging in our problem because the customer load demands show high non-stationarity. Significant variations in power consumption can be observed between peak times (afternoon and evening) and non-peak times (midnight and early in the morning). Thus, we need to estimate the transition matrices of the Markov Chain based on transition histories of all time steps. We derive the maximum likelihood estimation (MLE) of the transition matrix for one time step, and the same structure holds for all other MLE's in the rest of the time steps.

Denote by $\{p_{ij}|0 \leq i, j \leq m\}$ the entries of transition matrix and denote by X_t the state in step t, where p_{ij} is the transition probability from state i to state j and m is the number of states. For any pair of initial and final states x_1 and x_n, the likelihood is given by $L(p) = Pr(X_1 = x_1) \prod_{t=2}^{n} p_{x_{t-1}x_t}$. Define the transition

counts n_{ij} as the number of times that state i is followed by state j, then rewrite the maximum likelihood estimation problem as

$$\underset{p_{ij}}{\text{maximize}} \qquad Pr(X_1 = x_1) \prod_{i=1}^{m} \prod_{j=1}^{m} p_{ij}^{n_{ij}} \tag{2.1}$$

$$\text{subject to} \qquad \sum_{j} p_{ij} = 1, \quad i = 1, 2, \ldots, m \tag{2.2}$$

Notice that the optimal estimation $\hat{p}_{ij}$ that maximizes the log-likelihood $\log(L(p))$ also maximizes the likelihood function, where

$$\log(L(p)) = \log(Pr(X_1 = x_1)) + \sum_{i,j} n_{ij} \log(p_{ij}) \tag{2.3}$$

Then this convex optimization problem can be solved by introducing a new objective function $\mathbb{L}(p) = \log(L(p)) - \sum_{i}^{m} \lambda_i (\sum_j p_{ij} - 1)$, where $\lambda_1, \lambda_2, \ldots, \lambda_m$ are Lagrange multipliers. Taking into account both zero derivative conditions $\partial \mathbb{L}(p)/\partial p_{ij} = 0$ and the probability transition matrix constraints $\sum_j p_{ij} = 1$, for $i = 1, 2, \ldots, m$, we have $n_{ij}/\hat{p}_{ij} - \lambda i = 0$ and $\sum_{j=1}^{m} n_{ij}/\lambda_i = 1$ for all i. Thus, the MLE estimator of transition matrix is given by

$$\hat{p}_{ij} = n_{ij} / \sum_{j} n_{ij} \tag{2.4}$$

2.2.2 Prediction Performance Test against Real Measured Data

To validate the proposed Markov chain model with the transition matrix estimator derived above, we select some real measured data of 30 days from the huge data pool

of the Electric Reliability Council of Texas [53], in which both forecasted and actual power loads from clusters of households were recorded every 15 minutes for nearly 200 different locations. The Markov chain we adopt has 96 time steps corresponding to all 15 minute intervals in a day. Since in practice the number of states need to be finite, we first quantize data into a finite number of levels which are defined as states. The states are determined as follows: The entire range from minimum to maximum load demand is uniformly divided into consecutive intervals. The mean values of the uniformly divided intervals are adopted as the states of the Markov chain model. All data samples in a interval are then represented by the state value of that interval. Starting with the same initial distribution of the data, we generate load distributions using the derived MLE estimators of transition matrices for all time steps using half of the available data and thereafter find the load distribution in each time step of the Markov chain. The statistics of the distributions are then compared to that of the other half of the data.

Well matched results with average mean error below 1% and average standard deviation error below 10% can be observed. Figure 2.1 shows the mean values and standard deviations over time of the real load distribution and the predicted load distribution generated by a Markov chain with 6 states, with average mean error of 0.53% and average standard deviation error of 8.1%. Moreover, we also investigate the dependence of the performance of Markov chain models with different number of states. Figure 2.2 shows the variations of average errors in mean and standard deviation respectively, as the number of Markov chain state increases from 3 to 10. It can be seen that both errors decrease as the number of states increases. This matches our intuition because a bigger number of states means a smaller quantization interval size, including a smaller quantization error.

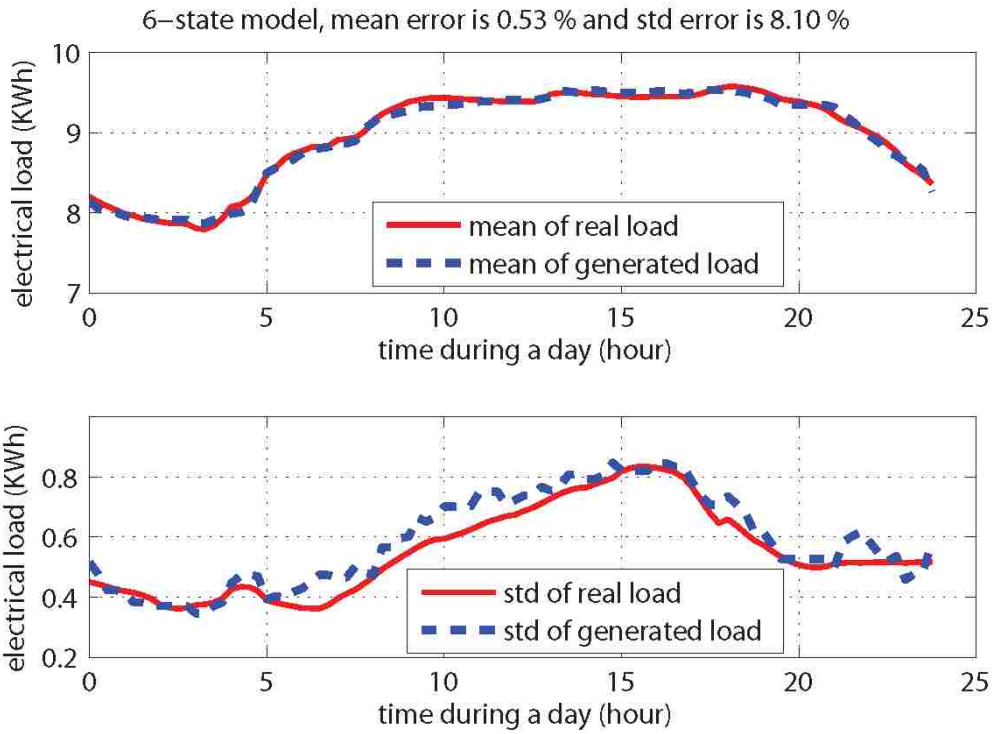

Figure 2.1: Mean values and standard deviations of real load and load predicted based on a 6-state Markov chain.

2.3 Linear Prediction Techniques for Customer Load Demand Modeling

Though the non-stationary Markov chain model provides a time varying linear description of the time transient property of the load demand probability distribution, it does not provide an easy way to predict the customer load demand directly based on the immediate load data history. Time series analysis techniques, however, provide good alternative approaches by which the future load demands are predicted by a linear function based on past load data record. For the prediction period with general length of several days, the standard autoregressive process model can not be

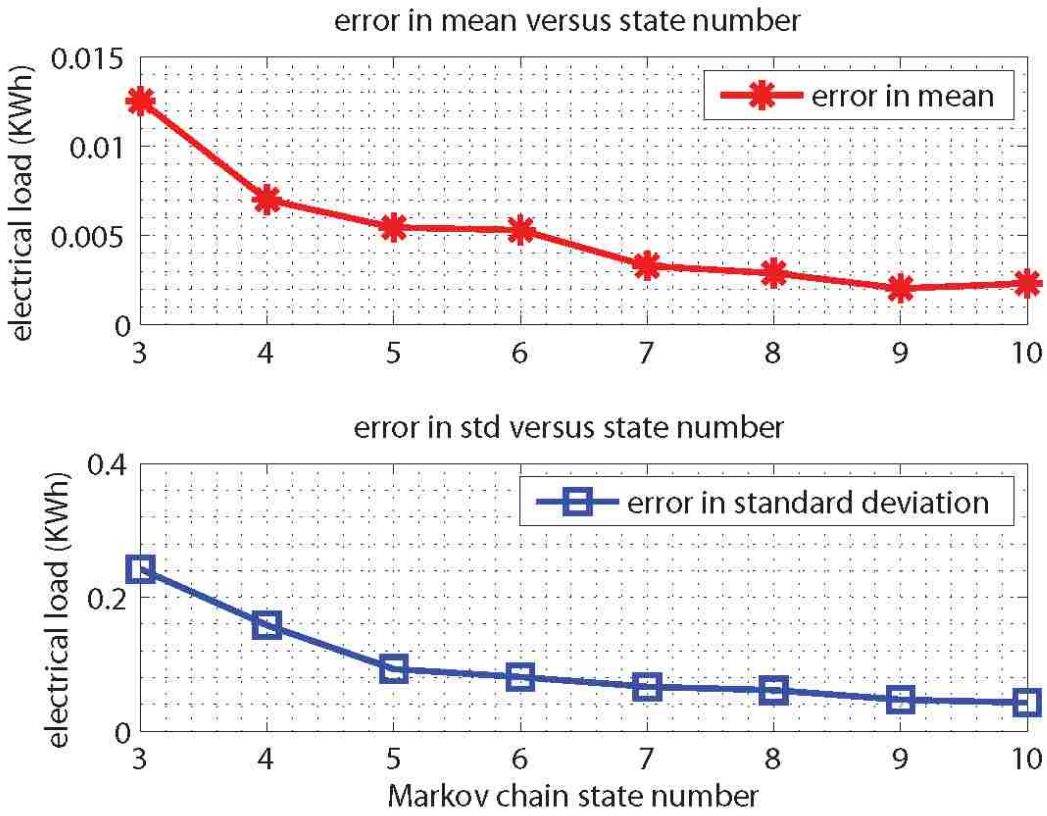

Figure 2.2: Errors in average mean value and standard deviation both decrease as the number of the Markov chain state increases.

directly used for load demand prediction because of the non-stationary load profile within the prediction period. Thus, we divide the entire load demand profile into consecutive short segments. The load data within the short segments are assumed to be stationary, which indicates piecewise stationarity over the entire prediction period. For each stationary segment, we apply a p-order AR model for the customer load demand modeling. Denote by $s(i)$ as the load demand at time instant i, then the AR model can be written as

$$s(i) = \sum_{j=1}^{p} \phi(j)s(i-j) + v(i) \tag{2.5}$$

where $\phi(1), \phi(2), \ldots, \phi(p)$ are the parameters of the model and $v(i)$ is the white noise.

To apply the AR model, there are two key issues to be addressed: determining the order p of the AR model and calculating the model coefficients $\phi(i)$, $i = 1, 2, \ldots, p$. To determine the model order, different approaches have been developed, such as Akaike information criterion (AIC), Bayesian information criterion (BIC) and partial autocorrelation function (PACF) approach [54].

2.3.1 Determining the AR Model based on PACF Analysis

In this work, we determine the model order based on the partial autocorrelation function (PACF) analysis. The PACF of lag k of a stationary sequence is defined as the autocorrelation of lag k with the linear dependence of lower order autocorrelation removed [55]. Mathematically, denoted by $\beta(k)$ the PACF of lag k, we have

$$\begin{aligned} &\beta(1) = Corr[s(i), s(i+1)], \quad \text{for} \quad k = 1 \\ &\beta(k) = Corr[s(i) - L_{s(i+1),\ldots,s(i+k-1)}(s(i)), \\ &s(i+k) - L_{s(i+1),\ldots,s(i+k-1)}(s(i+k))], \quad \text{for} \quad k \geq 2 \end{aligned} \tag{2.6}$$

where $L_{s(n_0),s(n_0+1),\ldots,s(n_1)}(s(m))$ denotes the projection of $s(m)$ onto the space spanned by $s(n_0), s(n_0+1), \ldots, s(n_1)$, or equivalently, denotes the best (in terms of minimizing MSE) linear estimate of $s(m)$ based on $s(n_0), s(n_0+1), \ldots, s(n_1)$. We may write

$$\hat{s}(m) = L_{s(n_0),s(n_0+1),\ldots,s(n_1)}(s(m)) = \sum_{i=n_0}^{n_1} a(i)s(i)$$

where $a(i)$, $i = s(n_0), s(n_0+1), \ldots, s(n_1)$ are determined by minimizing $E\{(s(m) - \hat{s}(m))^2\}$. Thus, the PACF at lag k may be interpreted as the correlation between $s(i)$ and $s(i+k)$ with the effect of the intermediate variables $s(i+1), \ldots, s(i+k-1)$ "filtered out". This is essential because usually the correlation of high order lags could be merely due to the propagation of the autocorrelation at lower order lags.

It has been shown in [56] that for an autoregressive process of order p, the PACF $\beta(k)$ will be nonzero for $k \leq p$ and zero for $k > p$. Thus, to fit the stationary data sequence in each segment, we determine the order of the AR model by analyzing the empirical PACF of the data with some approximate cutoff. We analyze the same set of real measured data as we used in the non-stationary Markoc chain approach. Figure 2.3 shows the empirical PACF of one stationary segment, with the threshold predefined (red dashed line), the reflection coefficients (The reflection coefficients constitute unbiased estimates of the partial correlation coefficients.) with lags greater than 4 are smaller than the threshold and are considered as zeros. Thus the AR model order for that stationary segment is 4.

2.3.2 Autoregressive Process Coefficient Estimation

Three methods of autoregressive-parameter estimation from data samples are usually considered in the literature, the least-square approach, the Yule-Walker approach and Burg's method [57]. The least-square approach and the Yule-Walker approach are pretty similar and differ only in the way how the autocovariance function is estimated. In least-square approach, unbiased estimate of the autocovariance function is used. While biased estimate of the autocovariance function is used in the Yule-Walker approach. The Levinson-Durbin algorithm provides a fast solution of a system of linear equations (e.g. Yule-Walker equation) containing a Toeplitz-style matrix. In contrast to the least-square and Yule-Walker method, which estimate the autore-

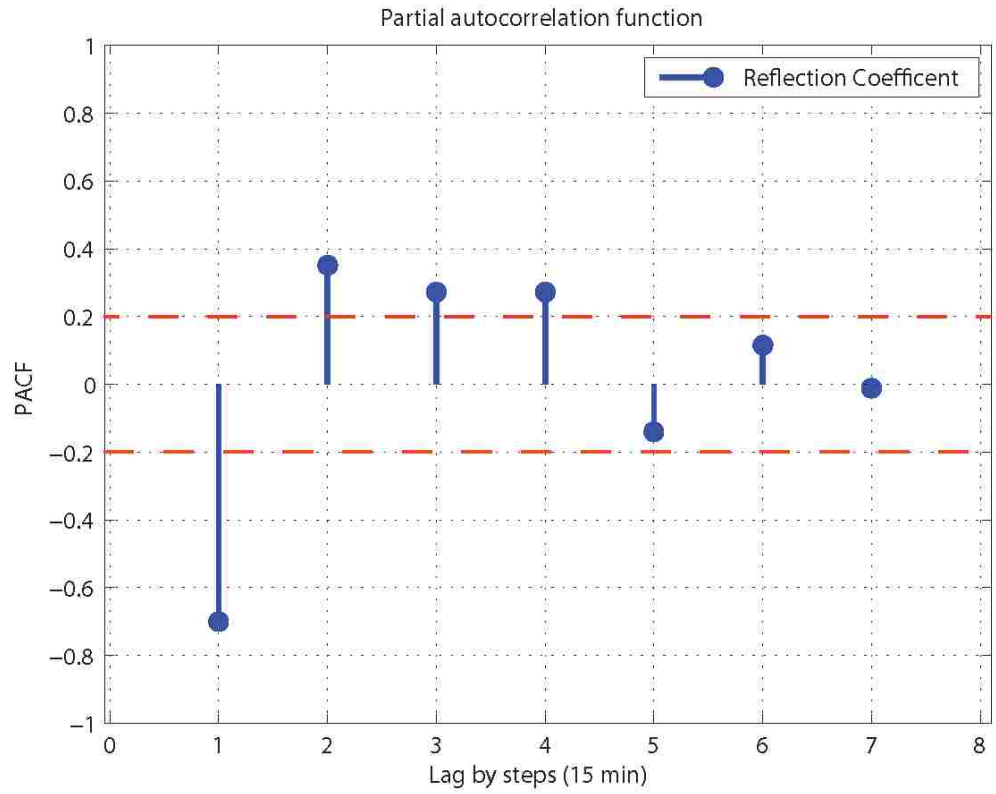

Figure 2.3: Empirical PACF of one stationary interval, with the threshold predefined (red dashed line), the reflection coefficients (PACF) with lags greater than 4 are smaller than the threshold value and are considered as zeros.

gressive parameters directly, Burg's method first estimates the reflection coefficients, which are defined as the last autoregressive-parameter estimate for each model order p. From these, the parameter estimates are determined using the Levinson-Durbin algorithm. To keep the discussion within the available space, we implement the Yule-Walker approach in this study.

We first make the data in each stationary segment zero-mean, by replacing $s(i)$ with $\tilde{s}(i) = s(i) - \bar{s}$, where $\bar{s} = \frac{1}{N}\sum_{i=1}^{N} s(i)$. We calculate the sample autocovariances $C(\tau) = \frac{1}{N}\sum_{i=1}^{N-\tau} \tilde{s}(i)\tilde{s}(i+\tau)$, for $\tau = 0, 1, \ldots, p$. Then the Yule-Walker equation can be written as follows, which can be solved using the Levinson-Durbin algorithm.

$$\begin{bmatrix} C(0) & \dots & C(p-1) \\ C(1) & \dots & C(p-2) \\ \vdots & \ddots & \vdots \\ C(p-1) & \dots & C(0) \end{bmatrix} \begin{bmatrix} \phi(1) \\ \phi(2) \\ \vdots \\ \phi(p) \end{bmatrix} = \begin{bmatrix} C(1) \\ C(2) \\ \vdots \\ C(p) \end{bmatrix} \tag{2.7}$$

We test the AR model with piecewise stationarity assumption against the same set of real measured data. The load profile is divided into 12 stationary intervals. For each segment, the AR order is determined based on the empirical PACF and the AR coefficients are determined based on part of the data. The predicted load profile is then compared to the rest of the data. Figure 2.4 shows the prediction performance of AR model with piecewise stationarity assumption. It can be observed that jumps exist between segments because different segments have different set of AR coefficients.

2.3.3 Predicted Load Profile Smoothing: a Local Stationarity Assumption

As mentioned in the prediction performance analysis for the AR model with piecewise stationarity assumption, there are abrupt jumps between consecutive segments in the predicted load profile because of the piecewise stationarity assumption. This is not a realistic assumption, because even though faster or slower changing rates can be observed in a real load profile, abrupt jumps in power consumption rarely appear. The smoothness of load profile becomes more apparent when more customers get involved. Because abrupt changes in individual consumption behavior may get averaged out in the total load consumption, when most of the customers have smooth load profiles at that time instant.

Thus, instead of assuming piecewise stationarity, we assume that the customer

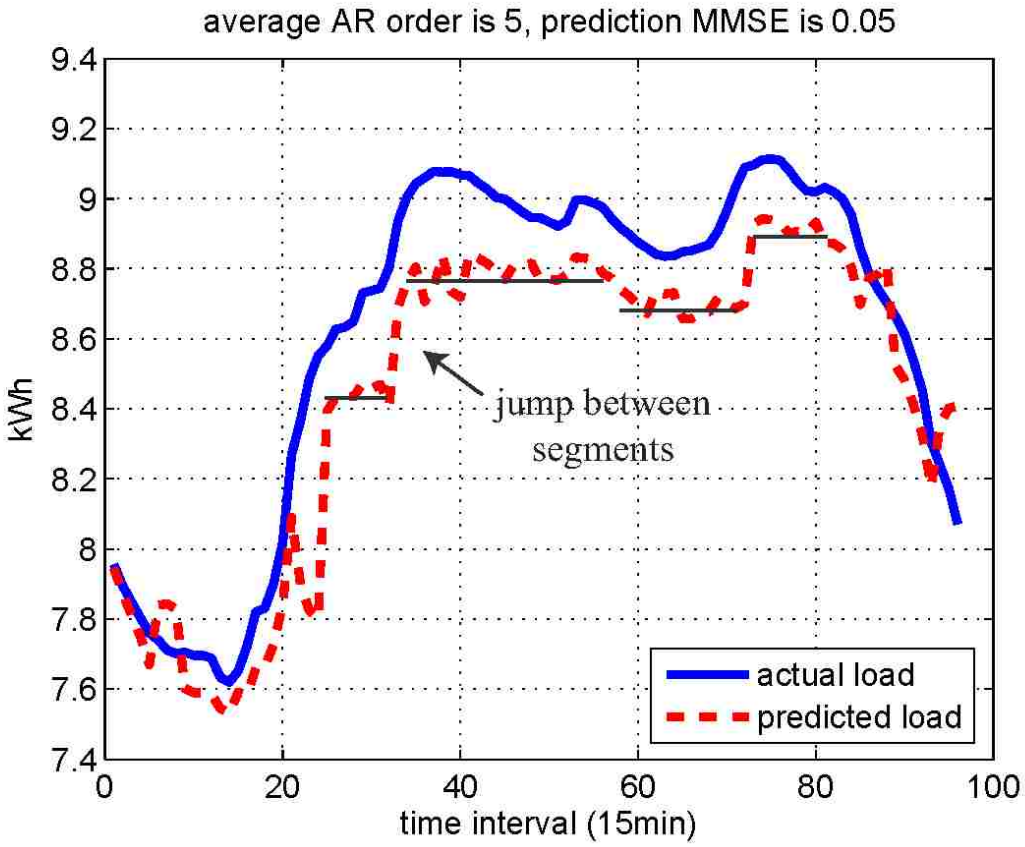

Figure 2.4: Prediction performance of AR model with piecewise stationarity assumption. It can be observed that jumps exist between segments because different AR coefficients are applied for different segments.

load profile is locally stationary, which means the data in a segment can be considered to be stationary as long as the segment is small enough. Based on this assumption, we can still apply the AR model for small time segments as we did before. The difference is that now a sliding window is applied to ensure a certain length of overlapping between consecutive segments. Figure 2.5 shows the prediction performance of the AR model with locally stationarity assumption, with much smoother predicted load profile and smaller prediction error. The sliding window size is the same as the segment length and the sliding step size is one forth of the segment length.

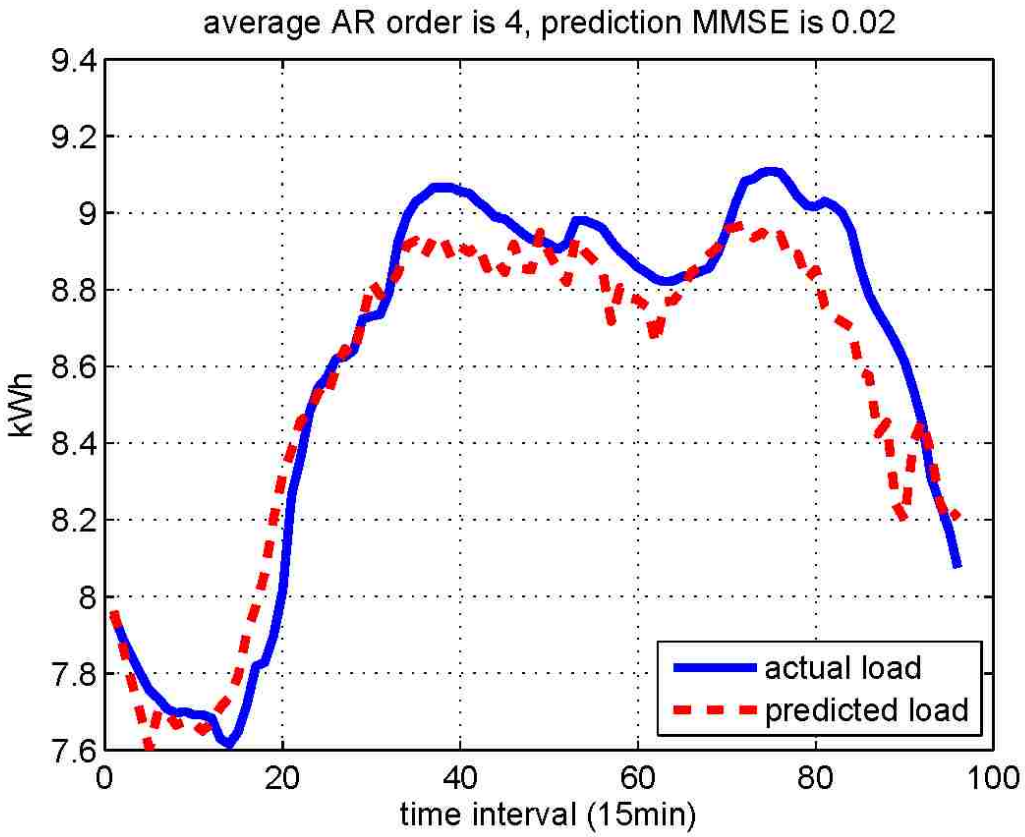

Figure 2.5: Prediction performance of AR model with locally stationarity assumption. The sliding window size is the same as the segment length and the sliding step size is 4.

2.4 Time Varying Autoregressive (TVAR) Process for Customer Load Demand Modeling

Though the prediction performance of the standard AR model is acceptable, dividing the load profile into segments, both under piecewise and locally stationarity assumptions, may degrade the prediction efficiency especially when the size of the segment need to be sufficiently small to guarantee the stationarity. For a prediction period with general length of several days, dividing the prediction period into a large

number of segments might not be efficient enough for real time operation in practice. However, close observation on the load profiles over consecutive days shows that the power consumption pattern over consecutive days have similar patterns (load variation profiles). Thus, we can assume that the load demands are cyclo-stationary with cyclic period of one day. Under this assumption, we can develop a prediction model based on the load data in one day, and the same model parameters can be adopted for load prediction in the following days. In this way, load prediction can be implemented in a very efficient way even for relatively long prediction periods. As the load profile within a day is non-stationary, we generalize the AR model by allowing the AR coefficients to be time variant, which leads to a TVAR process [54, 58]. Denote by $s(i)$ the load demand at time instant i, we assume a p-order TVAR model, which can be written as

$$s(i) = \sum_{j=1}^{p} \phi_j(i) s(i-j) + v(i) \tag{2.8}$$

where $\phi_j(i)$, $j = 1, 2, \ldots, p$ are the coefficients of the model, which are functions of both lag j and time i. $v(i)$ is the zero-mean white Gaussian noise. To estimate these time variant coefficients, we can further approximate each coefficient by a weighted combination of a set of $q+1$ independent "basis" functions $f_n(i)$, $n = 0, 1, \ldots, q$. Thus we have $\phi_j(i) = \sum_{n=0}^{q} \phi_{jn} f_n(i)$, where ϕ_{jn}'s are time invariant. Thus, we can rewrite the TVAR process as

$$s(i) = \sum_{j=1}^{p} (\sum_{n=0}^{q} \phi_{jn} f_n(i)) s(i-j) + v(i) \tag{2.9}$$

$$= \left[S_{i-1}^T, S_{i-2}^T, \ldots, S_{i-p}^T\right] \cdot \Theta + v(i) \tag{2.10}$$

where

$$S_i = [f_0(i)s(i), f_1(i)s(i), \ldots, f_q(i)s(i)]^T, \tag{2.11}$$

$$\Theta = [\phi_{10}, \ldots, \phi_{1q}, \phi_{20}, \ldots, \phi_{2q}, \ldots, \phi_{p0}, \ldots, \phi_{pq}]. \tag{2.12}$$

Now we can see the advantage of the "basis" function parameterizations: A linear non-stationary problem becomes a linear time-invariant problem by replacing a scalar process with a vector one [58]. To estimate the time invariant coefficients Θ that gives the optimal prediction $\hat{s}(i) = \left[S_{i-1}^T, S_{i-2}^T, \ldots, S_{i-p}^T\right] \cdot \Theta$, it is therefore meaningful to minimize the variance of the prediction error $s(i) - \hat{s}(i)$. We therefore obtain the optimum vector Θ as the solution of an equation of the Yule-Walker type

$$E\left(\begin{bmatrix} S_{i-1} \\ S_{i-2} \\ \vdots \\ S_{i-p} \end{bmatrix} \cdot \left[S_{i-1}^T, \ldots, S_{i-p}^T\right]\right) \cdot \Theta = E\left(\begin{bmatrix} S_{i-1} \\ S_{i-2} \\ \vdots \\ S_{i-p} \end{bmatrix} \cdot s(i)\right) \tag{2.13}$$

Note that in the TVAR process modeling, in addition to time variant coefficient estimation, there are other key issues to be addressed such as determining the order of the model and the selection of the set of basis functions [58]. To determine the optimal order p of the TVAR process that describes the behavior of the time series the best, the PACF analysis approach mentioned before can still be adopted. However, different from the PACF for stationary data sequence, the PACF for a non-stationary data sequence is not only a function of lag k, but also a function of time instant t. Let us denote the PACF of a non-stationary data sequence by $\beta(t, t-k)$. From the results in [55], it follows that a process $s(i)$ is autoregressive of order p if and only if its PACF $\beta(\cdot, \cdot)$ satisfies the following two conditions:

(1) $$\forall t \in \mathbb{Z} \text{ and } \forall k > p, \beta(t, t-k) = 0, \tag{2.14}$$

(2) $$\forall k \leq p, \exists t \in \mathbb{Z} \text{ s.t. } \beta(t, t-k) \neq 0. \tag{2.15}$$

According to [55], we can estimate the TVAR order of the entire load profile by averaging the AR orders of stationary segments generated by a sliding window. For another issue of selecting the basis function set, there are also a lot of choices. The most common ones include orthogonal polynomial functions, trigonometric functions, non-periodic Fourier basis and so on. Considering the cyclostationarity of the load data, we select the trigonometric functions. Note that the lowest frequency component of the basis function set should have a frequency that is of an integer multiple of the cyclic frequency of the load profile, in order to catch the periodicity of the load profile.

In the following, we test the proposed TVAR model against the same set of real measured data. In our simulation, the order of the applied TVAR model is 4. In the trigonometric basis function set, we set the lowest frequency component to be 6 times of the cyclic frequency. The expectation in the coefficient estimation are replaced by the sample means. From the 30 days of data, part of the data are used for training, and then the obtained TVAR model is test against the rest of the data. The top plot in Figure 2.6 shows the comparison between the actual and the predicted load profiles, with first 5 days' data as the training data. The bottom plot shows that the prediction MSE decreases as the length of training increase from 2 to 6 days.

It is worth pointing out that the prediction under the cyclostationarity assumption gradually becomes off from the real measured data when the prediction period becomes too long, i.e. over months or seasons. This is because customer power consumption pattern do change from season to season. Note that verifying the cyclostationarity of a data sequence in a meaningful way is not an easy topic and out of

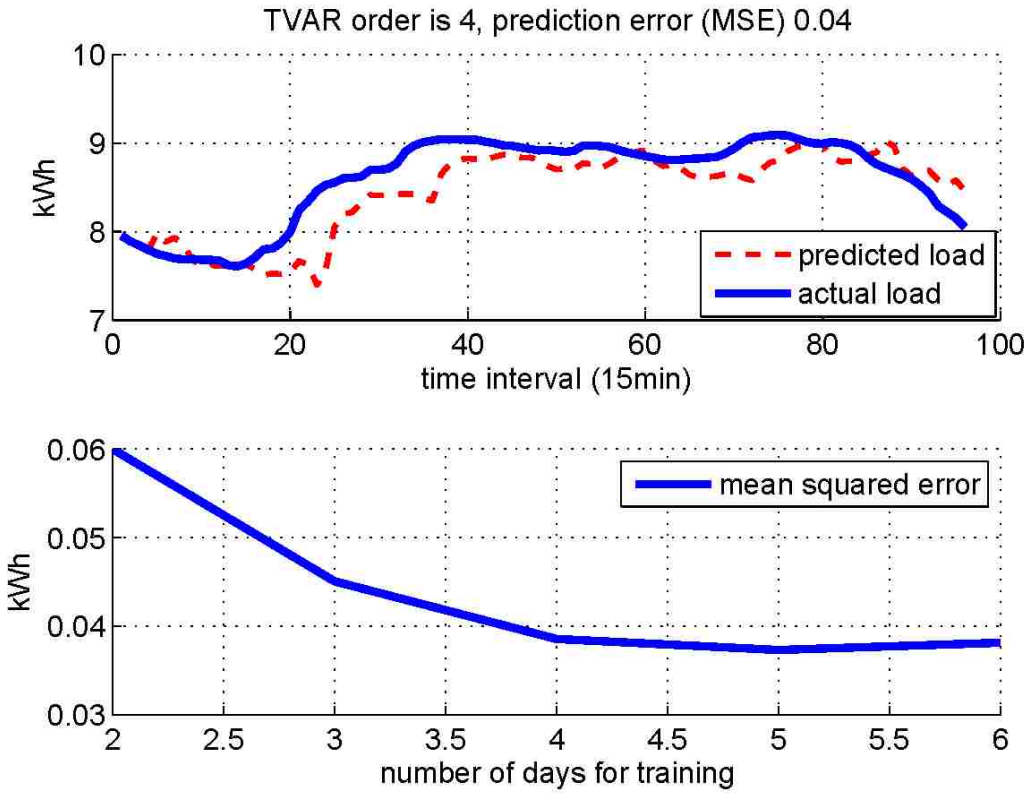

Figure 2.6: The top plot shows the prediction performance of the 4-order TVAR model, with the data of the first 5 days as training data. The bottom plot shows prediction error (MSE) decreases as the length of training increases from 2 to 6 days.

the scope of this work, since huge amount of other information is required to go with it. Interesting readers are referred to [59] and related references there. Hence even though the TVAR model increases the prediction efficiency, updates on the model coefficients are still necessary periodically (say, monthly) in practice.

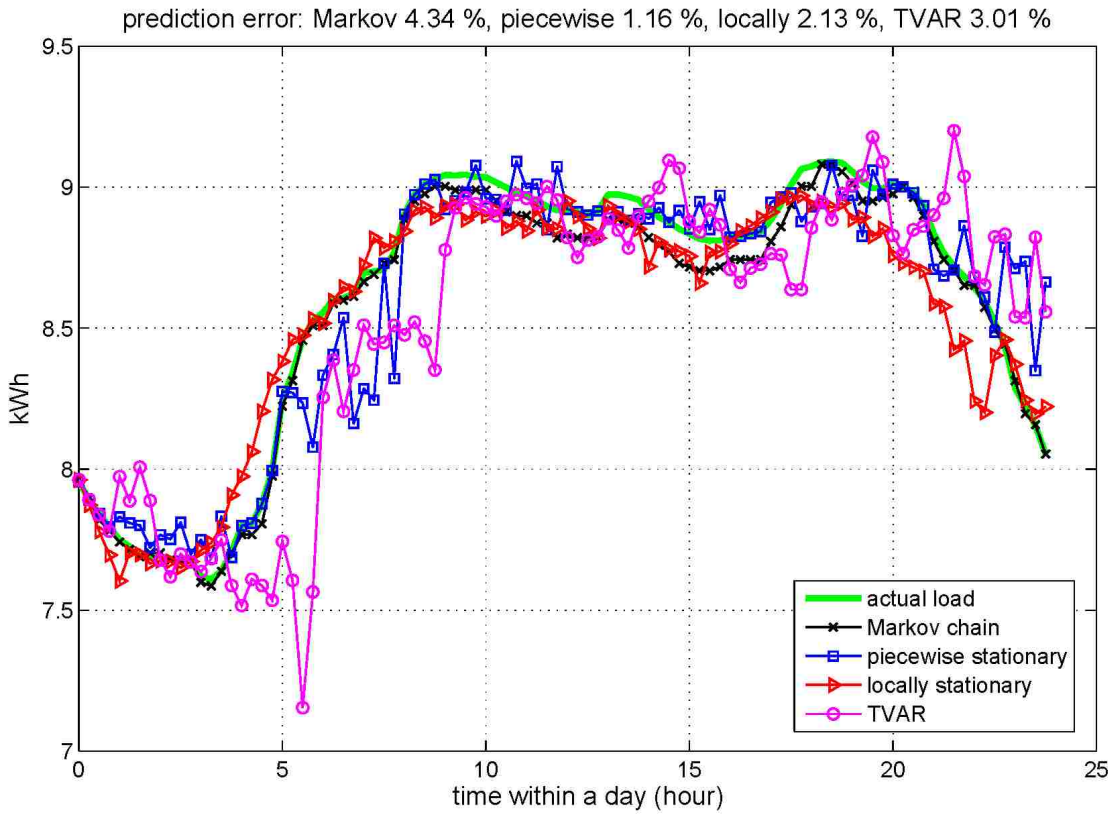

Figure 2.7: Prediction performance comparison among different modeling approaches. 20% of the selected data is used for training, the predicted load profiles are compared to the average of the rest 80% of the data.

2.5 Prediction Performance Comparison among Different Modeling Approaches

We compared the prediction performances of different modeling approaches we proposed, with the same simulation setup. 20% of the selected data is used for training, the predicted load profiles are compared to the average of the rest 80% of the data. The prediction performances of different modeling approaches are shown in Fig. 2.7. The advantages and disadvantages of all approaches are summarized in Fig. 2.8.

	Advantages	Disadvantages
Markov Chain	Estimate the non-stationary load demand distributions	Not enough emphasis on the immediate load data history
Piecewise stationary	Linear prediction based on the immediate load data history	Abrupt jumps between segments in the predicted load profile
Locally stationary	Smoother predicted load profile because of the introduced sliding window.	Degrade the prediction efficiency especially for long prediction period
TVAR	Significantly increase the prediction efficiency (less coefficient estimations)	Day of the week, monthly and seasonal effect should be considered

Figure 2.8: Advantages and disadvantages of different modeling approaches.

Based on the analyses on different load modeling and prediction schemes presented above, we can conclude that both non-stationary Markov chain and the time series analysis techniques provide good modeling of the time transient features of customer load demand. The transition matrix estimation in the former approach describes how the distribution of load demands evolves over time, the latter approach provides linear prediction schemes which predict future load based on the immediate load history. Among the three linear prediction models we proposed, though the model with locally stationarity assumption improve the performance compared to the model with piecewise stationarity assumption by smoothing the predicted load profile, both AR based model have the disadvantage of degrading the prediction efficiency because the coefficient estimation is implemented for each short stationary segment. The TVAR based prediction model is based on the cyclo-stationarity of

load profiles, and overcomes this disadvantage while guaranteeing good prediction performance.

2.6 Conclusion

In this chapter, we proposed two types of approaches to model the uncertainty in customer load demand. The first approach was based on a first order non-stationary Markov chain. A maximum likelihood estimator (MLE) was derived to estimate the time variant transition matrix of the Markov chain. The second approach was based on time series analysis techniques. We presented linear prediction models such as standard autoregressive (AR) process and time varying autoregressive (TVAR) process, according to different assumptions on the stationarity of customer load profile: piecewise stationarity, local stationarity and cyclo-stationarity. Prediction performances of different models were analyzed and compared, advantages and disadvantages were discussed.

Chapter 3

Price-based Demand Response Scheme Design in Smart Grid

3.1 Introduction

In most current electricity markets, fixed pricing schemes with constant rates are being widely used. Customers face retail electricity prices that are flat over months or even years [23]. A problem with fixed pricing schemes is the disconnection between short-term marginal electricity production costs and retail rates paid by customers, which leads to inefficient overall resource usage. Due to lack of information on generation costs, electricity consumption behavior of customers may not adjust to supply-side conditions. Thus fixed constant pricing results in suboptimal customer behavior as well as higher electricity costs than they would otherwise be in an optimally efficient system [24].

There is a growing consensus that Demand Response (DR) can play an important role in market design [25]. Lack of DR has been shown to be a major contributing factor for energy-market meltdowns [26]. In [23], for example, 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." DR not only reduces the capacity investments in peak generation units to serve occasional heightened demand, but also provides short-term reliability benefits as it can offer load relief to resolve system and local capacity constraints. There are two basic demand response options: Price-based demand response and incentive-based demand response. Price-based demand response includes real-time pricing (RTP), critical-peak pricing (CPP), and time-of-use (TOU) rates. Customers can respond to the price structure with changes in energy use, reducing their electricity bills if they adjust the timing of their electricity usage to take advantage of lower-priced periods and avoid consuming when prices are higher [23]. Incentive-based demand response schemes pay participants to reduce their loads at times requested by the program sponsor, triggered either by a grid reliability problem or high electricity prices. DR programs typically specify a method for establishing customers' baseline energy consumption level below which demand reductions are not allowed. In power systems, the energy requests that customers send to utility consist of two parts: nonflexible load request and flexible load request [27]. The nonflexible part is the minimum amount of energy that utility needs to provide at a specific time. The flexible part can be reallocated over time according to a certain load management strategy. For any load management strategy there are two common primary goals: peak load shaving and load profile flattening. Under real-time pricing, the electricity price is determined by real time load information.

In this chapter, a block scheduling model of load management is presented for price-based demand response scheduling. In this model, the size of the time block is set to be small enough so that all load shifting within the time block can be considered as cost free and acceptable to customers. The solution to this block processing problem can then be the basis for implementations of arbitrarily long scheduling

periods. Two types of real-time pricing schemes, linear pricing and threshold pricing, are discussed in this paper. We consider optimal demand-response when customers cooperate as a group as well as when each customer is only interested in minimizing its own cost. Naturally these two scenarios, as shown to lead to centralized and distributed optimizations.

The rest of this chapter is organized as follows: In Section 3.2, the system model and the problem formulation for block scheduling are presented. In order to minimize the Utility's cost, a convex optimization problem is formulated and solved analytically in section 3.3. A searching method is presented to find the water levels of the water-filling solutions numerically in section 3.4. In section 3.5, the original problem is decoupled and the customer-wise power allocation is optimized by a two dimensional water-filling solution. Performance analysis and comparison with simulation results are also presented. The conclusions from this chapter are given in section 3.6.

3.2 Problem Formation

We assume an electricity market consisting of one electrical Utility and K customers. A block processing model is adopted here in which load demands are scheduled in a periodic block-by-block manner. Each block consists of I time intervals and the size of each time interval is T hours. The interaction model between the Utility and customers can be described as follows [60–62]: At the beginning of each time block, all customers submit their *predicted load demands* of the current time block to the Utility based on their energy requirements and electricity pricing information. The predicted load demands from customers consist of two parts: nonflexible load demands and flexible load demands. The nonflexible load demands are the *basic* energy requirements of customers, which specify how much electrical energy is needed during each time interval of a time block. We denote by $l_{i,k}^N$ and $\tilde{l}_{i,k}^F$ $(i = 1, 2, \ldots, I,$

$k = 1, 2, \ldots, K$) the nonflexible and flexible load demands from customer k in time interval i, denote by $l_i^N = \sum_{k=1}^{K} l_{i,k}^N$ and $\tilde{l}_i^F = \sum_{k=1}^{K} \tilde{l}_{i,k}^F$ the total nonflexible and flexible load demands over all customers in time interval i, similarly $l_N = \sum_{i=1}^{I} l_i^N$ and $\tilde{l}_F = \sum_{i=1}^{I} \tilde{l}_i^F$ denote the total nonflexible and flexible load demands during that time block. We assume that the Utility guarantees supporting all nonflexible load demands during each specified time interval. Thus, in any time interval the sum of nonflexible load demands over all customers is assumed to be no greater than the generation capacity. On the other hand we assume that no customer cheats on its nonflexible load demand, e.g. declaring more nonflexible load demand than its actual basic requirement. The demand management scheme design with self-oriented customers, meaning that customers might cheat to increase their own profits, is investigated in our early work [31] and [14]. From the beginning of a time block all customers want their flexible load demands be supported as early as possible. Energy supply from the Utility during later time intervals induces a delay cost which is an increasing function of both the delay time and the amount of energy that has been delayed, due to the dissatisfaction of customers. The delay cost defined above is important for the demand management scheme design as it addresses the requirement for a timely energy supply [63].

With the household load demand prediction technique developed in the previous section, the Utility and customers interact as follows: Customers submit their predicted load demands (both flexible and nonflexible loads) to the Utility by using the linear prediction techniques. Upon receiving customers' load demands, the Utility checks the generation capacity constraint and determines an optimal generation profile that minimizes its cost over time. On the generation side, the objective of the Utility is to minimize its cost. On the consumption side, the objective of customers is to maximize the profit, in terms of individual profits and/or the social welfare defined for the entire collection of customers. The whole interaction procedure between the Utility and customers is illustrated in Fig. 3.1.

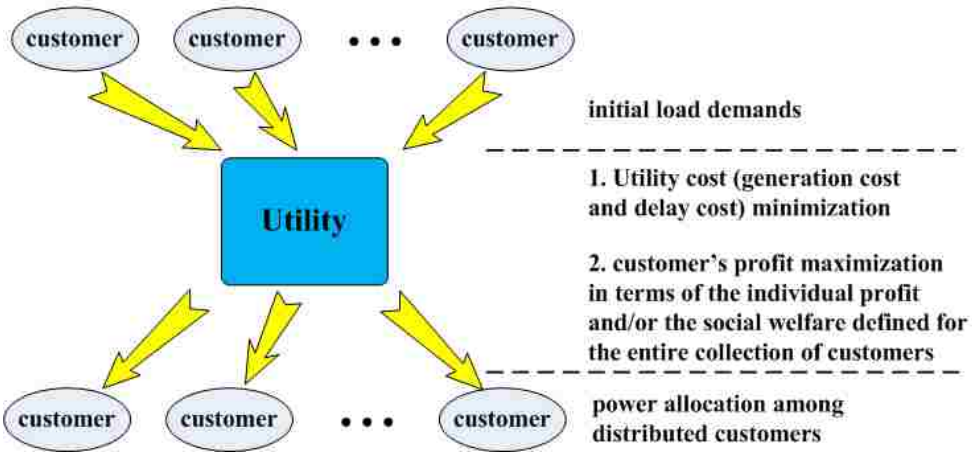

Figure 3.1: Interaction model between the electrical Utility and customers.

3.3 Utility Cost Minimization: Load Demand Scheduling over Time

As discussed in the problem formulation section, upon receiving the predicted load demands from all customers, the Utility can apply the DR scheduling for the purposes of peaking shaving and load profile flattening. In this section, we formulate and solve the problem of minimizing the Utility cost by scheduling the flexible load demands over time intervals within the processing block, assuming specific generation cost and delay cost forms. The amount of flexible load demand after load reallocation in time interval i is denoted by x_i and we assume that the Utility provides constant power of $\frac{l_i^N + x_i}{T}$ within each time interval. We assume that in time interval i, the generation cost per unit energy (in monetary measure) is a linear function of the total load demands in that time interval, say, $\alpha(l_i^N + x_i)$, where α is a positive scaling factor [30]. Thus the generation cost for time interval i and for the entire time block are given by $C_{g,i} = \alpha(l_i^N + x_i)^2$ and $C_g = \sum_{i=1}^{I} C_{g,i}$ respectively. Moreover, we assume that if x_i amount of load demands have been delayed by i time intervals, the associated delay

cost is given by $C_{d,i} = iTx_i\gamma^{-(I-i)T}$, where γ is the positive delay cost scaling factor. Thus the total delay cost for a processing block is $C_d = \sum_{i=1}^{I} C_{d,i}$. It is worth pointing out that, though in practice the cost terms could have different forms according to specific scenarios, the cost function defined in this paper is actually independent from the solution approach developed in this paper. When the total load demands (the nonflexible plus the flexible) in a time block is greater than the total energy that can be generated during that time block, no optimal reallocation solution exists unless the Utility cuts down the flexible load demands. For fairness, the following strategy is adopted: If the total amount of load demand is greater than the total generation capability, the Utility will cut every customer's flexible load demand by the same proportion to keep the total load demand equal to the generation capability, i.e, if $\tilde{l}_F + l_N > IL_M$, where L_M is the constant generation capacity of Utility in each time interval. The new flexible load demand is $l_F = \sum_{k=1}^{K} l_k^F$ and $l_k^F = \tilde{l}_k^F - \beta(\tilde{l}_k^F + l_k^N)$ where $\beta = \frac{\tilde{l}_F + l_N - IL_M}{\tilde{l}_F + l_N}$ for all k. Based on this centralized load cutting scheme, we may, without loss of generality, assume that the load demand in any time interval never exceed the generation capacity and thus drop the generation capacity constraint. Denoting by weighted sum $C = C_g + \delta C_d$ the total cost of Utility, where δ is the weight coefficient for delay cost, we have the following optimization problem

$$
\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & C(\mathbf{x}) = \alpha \sum_{i=1}^{I} (x_i + l_i^N)^2 + \delta \sum_{i=1}^{I} (iTx_i\gamma^{-(I-i)T}) \\
\text{subject to} \quad & -x_i \leq 0, \quad i = 1, 2, \ldots, I, \\
& \sum_{i=1}^{I} x_i - l_F = 0.
\end{aligned}
$$

With the primal problem being convex, the optimal primal and dual solutions are achieved if and only if the following Karush-Kuhn-Tucker (KKT) conditions are held [64]:

$$\begin{aligned} \sum_{i=1}^{I} x_i^* - l_F \quad &= 0 \\ -x_i^* \quad &\leq 0, \forall i, \\ \lambda_i^* \quad &\geq 0, \forall i, \\ -\lambda_i^* x_i^* \quad &= 0, \forall i, \\ \frac{\partial L(\mathbf{x}^*, \lambda^*, v^*)}{\partial x_i^*} \quad &= 0, \forall i. \end{aligned} \tag{3.1}$$

Solving the KKT conditions above, the optimal solution can be written as

$$x_i^* = \begin{cases} 0 & \text{if } w^* < \hat{w}_i \\ w^* - \hat{w}_i & \text{if } w^* \geq \hat{w}_i, \end{cases} \tag{3.2}$$

where $\hat{w}_i = l_i^N + \frac{\delta}{2\alpha}(iT\gamma^{-(I-i)T})$ and w^* is the unique solution to

$$\sum_{i=1}^{I} \max(0, w^* - \hat{w}_i) = l_F, \tag{3.3}$$

Note that, the left hand side of (3.3) is a piecewise-linear increasing function of w^*, with breakpoints at $l_{i,k}^N$, ensuring the uniqueness of its solution. In general, there may not be a closed form solution for w^*, requiring numerical computation. For $\delta = 0$ (i.e., delay cost is completely ignored), the solution (3.2) reduces to

$$x_i^* = \begin{cases} 0 & \text{if } w^* < l_i^N \\ w^* - l_i^N & \text{if } w^* \geq l_i^N \end{cases}, \tag{3.4}$$

This solution structure (3.4) is well known in information theory and is referred to as the *water-filling solution* [65]: We can think of l_i^N as the height of the bottom

level at location i along the time axis within the processing block. Starting from zero, we allocate flexible loads to the location with the lowest nonflexible load. As flexible loads increase, some of them are put into locations with higher nonflexible loads. We continue to allocate flexible loads in this way until we have allocated all of l_F. At this time, the height of the flat flexible load level would be the solution w^* of (3.4). This process is similar to the way in which water distributes itself in a vessel. The depth of water at location i is then the optimal value x_i^*.

To better interpret the solutions above, we consider an electricity network during a time block of $I = 24$ intervals with $T = 1$ hour. The customers' load demands of different time intervals are generated according to different distributions corresponding to time dependent electrical energy consumption behavior. Given a set of initial load demands, the optimal allocation results for different δ values are shown in Fig. 3.2. It is seen that the solution (3.2) is slightly different from the water-filling result as there is no constant water level when delay cost is considered. This is because the allocation results x_i $(i = 1, 2, \ldots, I)$ is determined not only by the nonflexible load l_i^N, but also by another time interval dependent term $\frac{\delta}{2\alpha}(iT\gamma^{-(I-i)T})$. Indeed, solution (3.2) is of a water-filling like form if we interpret $\hat{w}_i = l_i^N + \frac{\delta}{2\alpha}(iT\gamma^{-(I-i)T})$ as the new modified *nonflexible load* in which $\frac{\delta}{2\alpha}(iT\gamma^{-(I-i)T})$ acts as an additional time related nonflexible load. The water level drops over time since later time intervals induce greater additional *nonflexible load demands*. It can be seen that the water level gets steeper as the delay cost weight δ increases. However, once the load demand of a time interval achieves L_M, no more load demands can be allocated to that interval. Thus after δ increases to a certain value, the optimal load profile becomes saturated (fixed). In this saturated profile, all except the last time interval with positive flexible load demands get L_M amount of flexible load demands. The generation cost also achieves its maximum value corresponding to the saturated profile.

Figure 3.3 shows the saturated load profile (with L_M normalized to 1) and gen-

eration costs (with the minimum normalized to 1) for different values of δ, given the average initial load demands. Note that the generation cost is nondecreasing over δ's, which is intuitively reasonable because the Utility has smaller and smaller flexibility on the scheduling operation as δ increases.

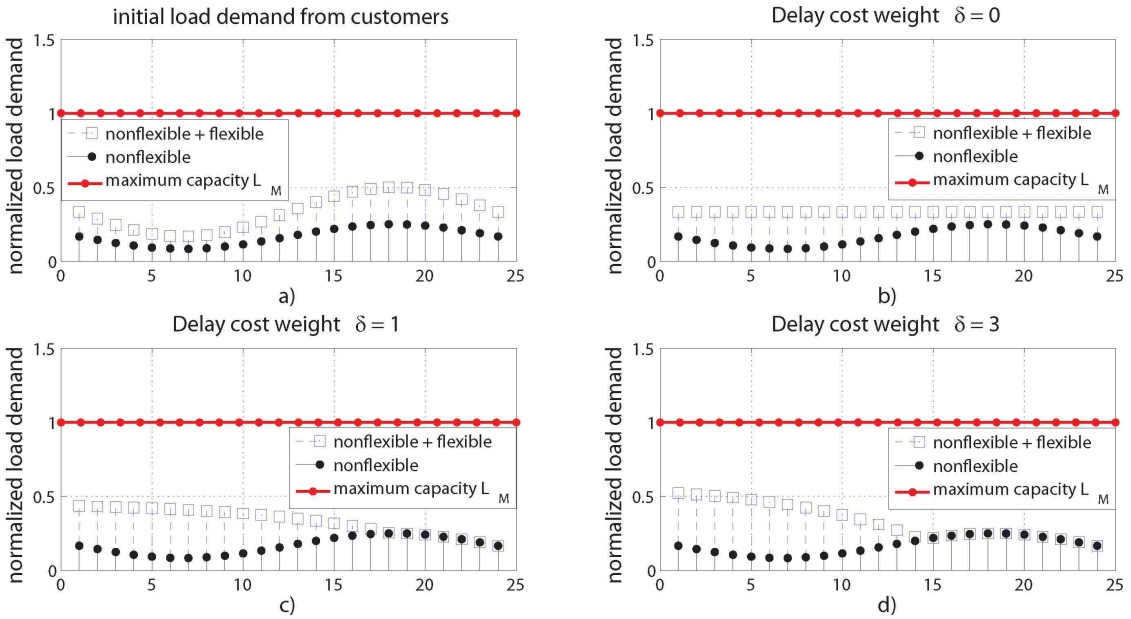

Figure 3.2: Optimal load profile comparison for different delay cost weights (δ's): a) Initial load demand from customers. b) Optimal load profile with no delay cost. c) Optimal load profile with delay cost weight $\delta = 1$. d) Optimal load profile with delay cost weight $\delta = 3$.

3.4 A Numerical Searching Method for Water-filling Solutions

As mentioned above, there is no closed form expression for the water level in the water-filling solution. In this section, we propose a searching method to find the water level. To keep the discussion general enough, we consider a threshold pricing

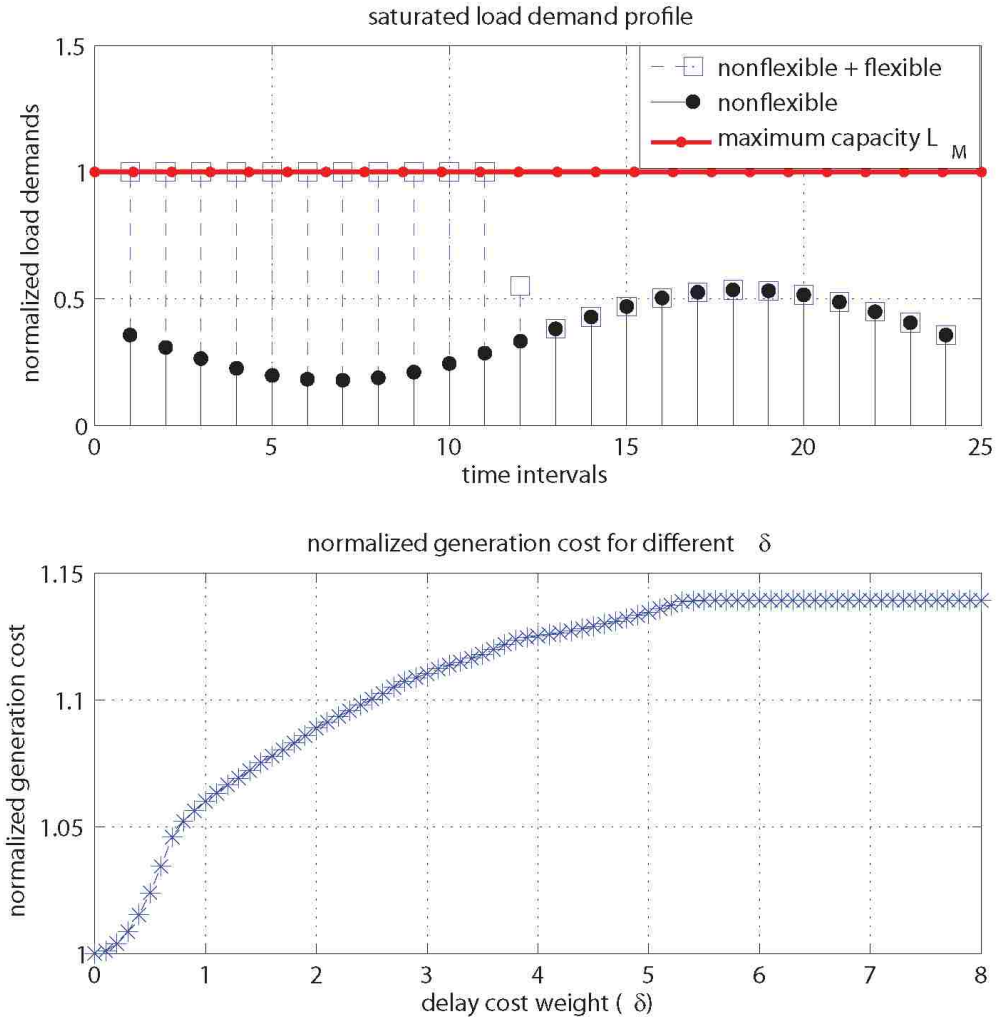

Figure 3.3: Saturated optimal load demand profile (upper) and generation cost for different δ's (lower)

scheme in this section. The linear pricing scheme we discussed earlier is actually a special case of the threshold pricing scheme with a threshold level of 0.

3.4.1 Threshold Pricing Scheme

Given initial load requests from customers, for a small enough load tuple Δl, all loads $\{l_i^F\}$'s and $\{l_i^N\}$'s can be represented as multiples of Δl. We label the m-th load tuple in time interval i by e_i^m, for $i = 1, 2, \ldots, I$ and $m = 1, 2, \ldots, M_i$, where $M_i = \frac{l_i^F + l_i^N}{\Delta l}$. In time interval i, there is a threshold L_i which can also be represented as multiples of Δl, say $L_i = \tilde{M}_i \Delta l$. We denote the price level of e_i^m by n_i^m. The price level n_i^m for tuple e_i^m is given by

$$n_i^m = \begin{cases} 0 & \text{if } m \leq \tilde{M}_i \\ m - \tilde{M}_i & \text{if } m > \tilde{M}_i \end{cases}, \tag{3.5}$$

$$= [m - \tilde{M}_i]^+ \quad . \tag{3.6}$$

Denote by $L_{\max}$ the maximum load capacity of Utility and $L_{\max} = M_{\max} \Delta l$. Then we have that $n_i^m \leq M_{\max}$ for $\forall i, m$, as shown in Fig. 3.4.

The threshold pricing scheme can be described as follows: In time interval i, a constant basic unit price P_0 (\$/kWh) applies for all e_i^m's below threshold L_i. The unit price for the m-th load tuple e_i^m in time interval i above threshold L_i is given by $P_i^m = P_0 + n_i^m \Delta P$, where ΔP (\$/kWh) is the increment in unit price. Denoted by x_i the flexible load request that Utility will schedule in time interval i, the consumption cost related to time interval i is given by $C_i = P_0(l_i^N + x_i)$, if $l_i^N + x_i \leq L_i$; and $C_i = P_0 L_i + \sum_{n_i^m=1}^{M_i - \tilde{M}_i} (P_0 + n_i^m \Delta P) \Delta l$, if $l_i^N + x_i > L_i$.

3.4.2 Water Level Searching Methods

We define the *vacancy* (shown in Fig. 3.4) value in time interval i as $v_i = M_i + 1$. Based on the observation that the only way of decreasing the generation cost is to

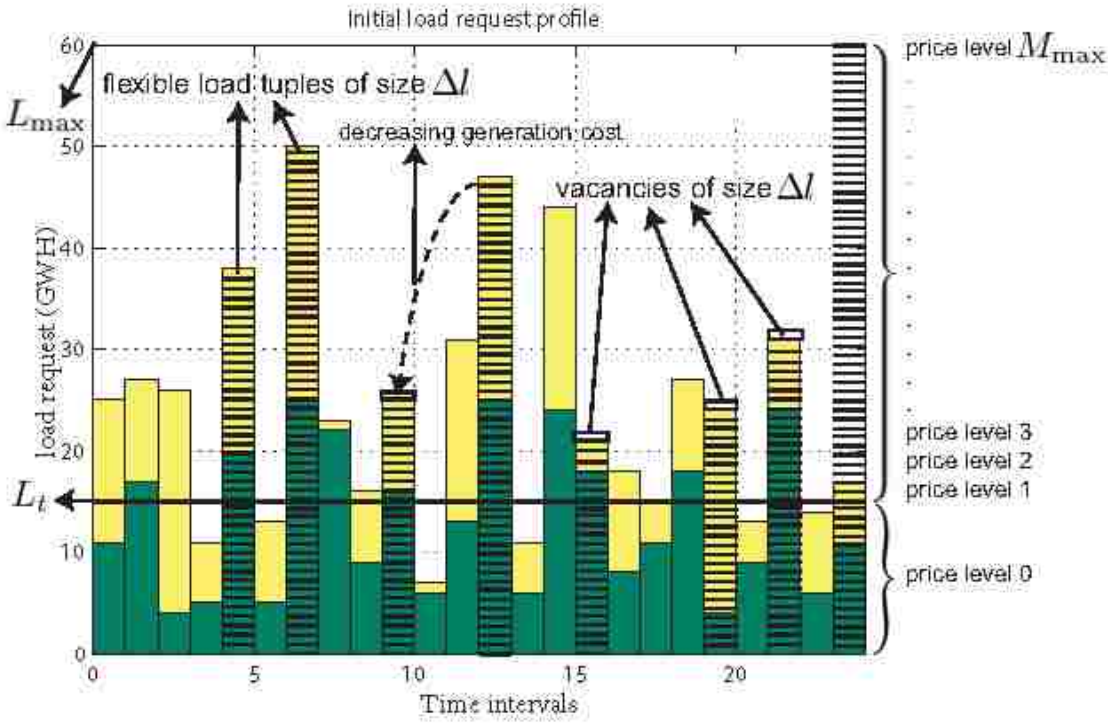

Figure 3.4: An illustration of threshold pricing scheme with discrete load tuples and vacancies. The threshold $L_{i,k}$ is set to be constant over all time intervals which could be dynamic in general.

shift some e_i^m's from higher price levels to vacancies with lower price levels, we have the following proposition:

Proposition 1. *In the block scheduling for threshold pricing scheme, the total consumption cost of all customers is minimized if and only if* $\max_i M_i \leq \min_i v_i$.

Proof. Necessity: Assume we have minimized the generation cost but there are some time-customer pairs i_1 and i_2 such that $M_{i_1} > v_{i_2}$, then by shifting the load tuple $e_{i_1}^{M_{i_1}}$ from price level M_{i_1} to the vacancy v_{i_2} we can further decrease the cost, this contradicts the minimum cost assumption.

Sufficiency: If the price cost function is not minimized, then there exists some load

shifting strategy that enable us to further decrease the generation cost. Thus there exists time-customer pairs i_3 and i_4 such that $M_{i_3} > v_{i_4}$. Thus if $\max_i M_i \leq \min_i v_i$ holds, there will be no load shifting strategy that could further decrease the cost function, meaning the current generation cost is the minimum. □

For the threshold pricing scheme, the optimal load profiles are of two categories according to whether the increment in unit price applies or not.

No increment in unit price applies

In the initial load profile, if all e_i^m's above the threshold can be allocated into vacancies below the threshold, then all the flexible loads in the optimal load profile will be in price level 0. All optimal load profiles that satisfy this property are considered as being optimal, meaning the optimal solution is not unique.

Increment in unit price applies

In the initial load profile, if the e_i^m's above the threshold are more than the vacancies below the threshold, then some flexible load tuples will cause price increments at some time intervals in the optimal load profile.

A slight variation of proposition 1 tells more about the optimal load profile in this case: Noticing that $\max_i M_i \leq \min_i v_i \Leftrightarrow \max_i v_i \leq \min_i v_i + 1$, the optimal load profile is flat in a Δl-flat sense. By "Δl-flat" we mean that $\max_{(i_1),(i_2)} |(l_{i_1}^N + x_{i_1}) - (l_{i_2}^N + x_{i_2})| \leq \Delta l$. As $\Delta l \to 0$, we have $\max_i v_i = \min_i v_i$. Thus, the optimal load profile again converges to a water-filling result. Hence, the optimization problem to minimize the Utility generation cost can be stated as follows: *Given initial load request information: price levels n_i^m's (M_i's) and vacancy levels v_i's for $t = 1, 2, \ldots, T$ with threshold level L_i, by doing a load reallocation which is also an updating process of n_i^m's and v_i's,*

we can minimize the total consumption cost of all time intervals if and only if the achieved load profile (possibly not unique) with M_i^*'s and v_i^*'s satisfy the optimization condition: $\max_t M_i^* \leq \min_i v_i^*$.

To find the optimal load profile, we may start from $\min_i v_i$ and search upward to $\max_i v_i$ until the testing level v_i^* satisfies the following conditions:

1. In the initial load request profile, the number of e_i^m's above the testing level v_i^* is strictly less than the number of all vacancies on and below testing level v_i^*.

2. In the initial load request profile, the number of e_i^m's on and above the testing level v_i^* is equal to or greater than the number of all vacancies below testing level v_i^*.

3.4.3 Simulation Results

The searching process and the termination condition described above gives the water level in the "Δl-flat" sense. In the following, we simulated the proposed load management strategy for an electric Utility with the threshold generation cost model during a period of $I = 24$ hours with each time interval $T = 1$ hour, and for a electricity market of 20 customes. Load tuple size is set to be $\Delta l = 1$ KWh. The Utility has a maximum capacity $L_{\max} = 60$ KWh. The threshold was set to be $L_i = 15$ KWh, $\forall t$. For each time interval, the flexible and nonflexible loads were generated according to uniform distributions $U(0, u_t)$ and u_t's were adjustable. As $L_{\max}$ is normalized to 1, all loads can be expressed as certain percentages of maximum Utility capacity.

Fig. 3.5 shows that in the initial load request profile, the number of flexible loads above the threshold ($L_i = 15$ KWh) is greater than the number of vacancies below the threshold. Thus, in the optimal load profile, extra increment in unit generation cost will apply. However, since the nonflexible loads are not high, the optimal load profile keep "Δl"-flat. Since the goal of DR is peak-load shaving and load-profile flattening,

we may set lower threshold L_i's for heavy load time intervals (for example, during day and evening hours) and higher threshold L_i's for lighter load time intervals (for example, during midnight hours) [66]. Such dynamic price-thresholding can naturally incentivize the customers to schedule their demand-responses in a way that will lead to peak-load shaving and load profile flattenning. An example of the dynamic threshold pricing is shown in Fig. 3.6. It can be observed that customers are encouraged to make use of off-peak time intervals. It is worth pointing out that though the profile in the right plot seems non-flat, it is actually flat in a price-level sense, as the surface of the flexible loads is in the same price level.

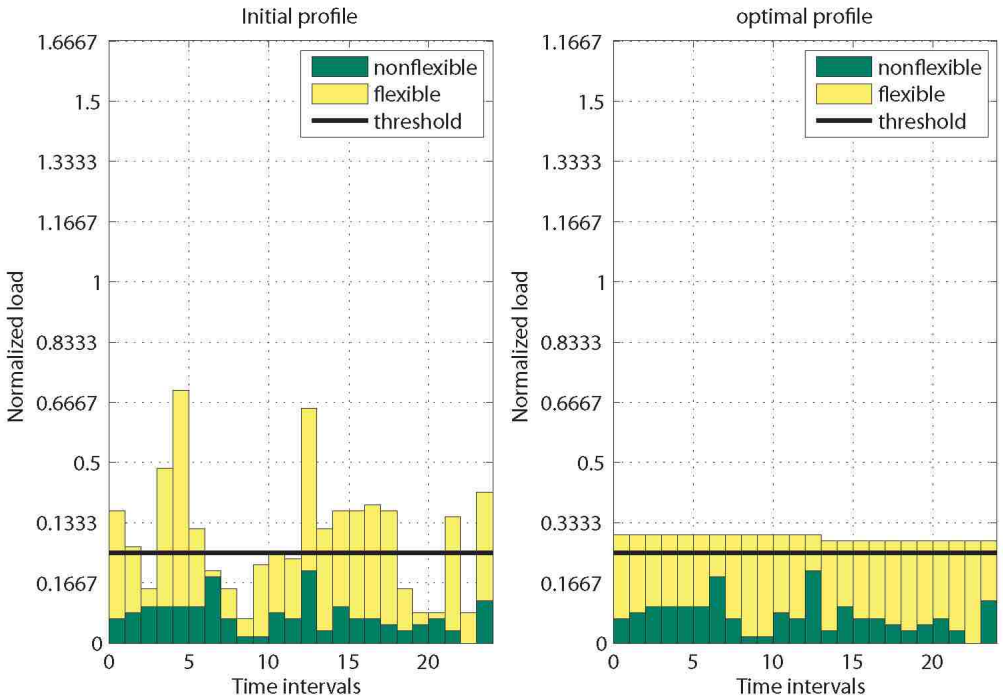

Figure 3.5: "Δl"-flat optimal profile with increment in unit generation cost.

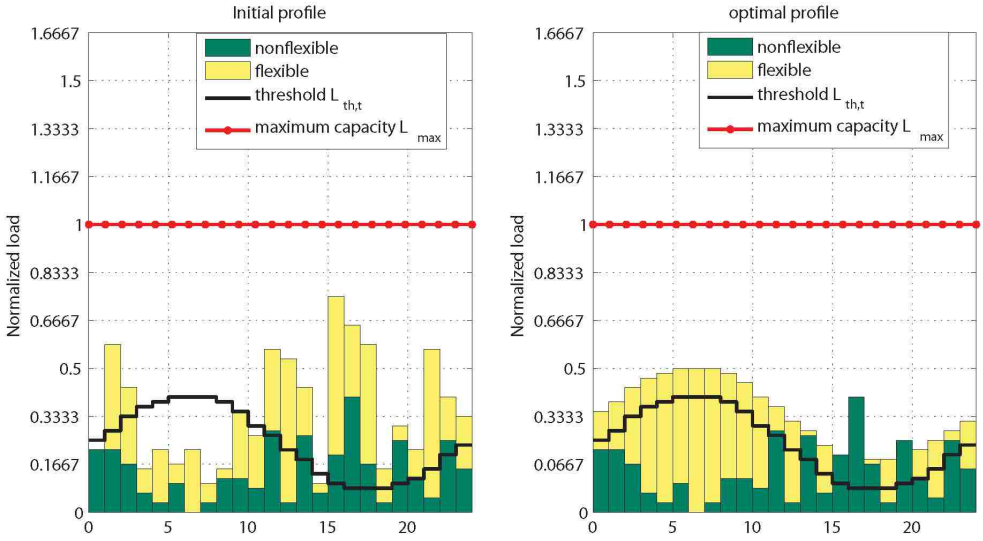

Figure 3.6: Optimal load profile with dynamic threshold pricing scheme. Though the profile in the right plot seems non-flat, it is actually flat in a price-level sense, as the surface of the flexible loads is in the same price level.

3.5 Customer Side Power Allocation: a Two-dimensional Water-filling Solution

The one-dimension water-filling solution presented above actually reveals a centralized optimization approach for power allocation among distributed customers. Consider the following optimization problem with K customers in the electricity market,

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & C(\mathbf{x}) = \alpha \sum_{k=1}^{K} \sum_{i=1}^{I} (x_{i,k} + l_{i,k}^N)^2 + \\ & \delta \sum_{k=1}^{K} \sum_{i=1}^{I} (iT x_{i,k} \gamma^{-(I-i)T}) \\ \text{subject to} \quad & -x_{i,k} \leq 0, \quad i = 1, 2, \ldots, I, k = 1, 2, \ldots, K, \\ & \sum_{k=1}^{K} \sum_{i=1}^{I} x_{i,k} - l_F = 0. \end{aligned} \tag{3.7}$$

It can be observed that this optimization problem is just an extension of the previous one but with one more dimension of customers. Again, by solving the KKT conditions, solutions with similar structures can be derived. In the special case with $\delta = 0$, the two-dimensional water-filling solution is given by (3.8), as shown in Fig. 3.7.

$$\begin{aligned} x_{i,k}^* &= \begin{cases} 0 & \text{if } w^* < l_{i,k}^N \\ w^* - l_{i,k}^N & \text{if } w^* \geq l_{i,k}^N \end{cases}, \\ &= [w^* - l_{i,k}^N]^+ \quad . \end{aligned} \tag{3.8}$$

If we sum the two-dimension water-filling profile over all customers, we get a one-dimensional profile, as shown in Fig. 3.8. Comparing Fig. 3.2(b) and Fig. 3.8, it can be concluded that the minimum Utility cost specified by the two-dimension water-filling solution is higher than the minimum Utility cost specified by the one-dimensional water-filling solution. Because in the former scenario the water level is strictly flat, while in the later scenario the water level of the one dimensional profile is not completely flat. This deviation between one-dimension and two-dimension water-filling solutions actually reveals how the Utility cost minimization can be affected by the customer clustering.

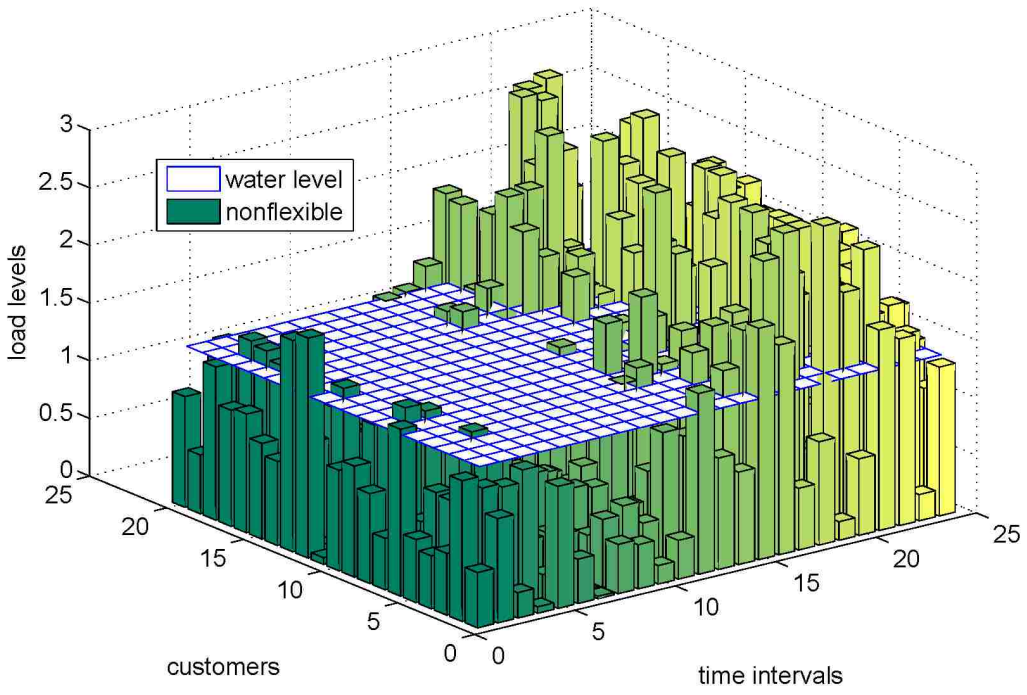

Figure 3.7: Two-dimensional water-filling solution that indicates how loads from different customers are scheduled over the processing time block.

3.5.1 Customer Clustering Effect

It is worth pointing out that if we switch the order of minimization and summation in the optimization problem (3.7), meaning that either we do the minimization for each customer first and then take the sum of them, or we just minimize the total cost of all customers. These two objective functions lead to different optimal solutions. The optimization problem (3.7) actually corresponds to the scenario in which all K customers are considered in one cluster. Denote by C^* the minimum cost corresponding to the scenario that all K customers are in one cluster, which is defined as follows (For simplicity, in this section we ignore the delay cost.),

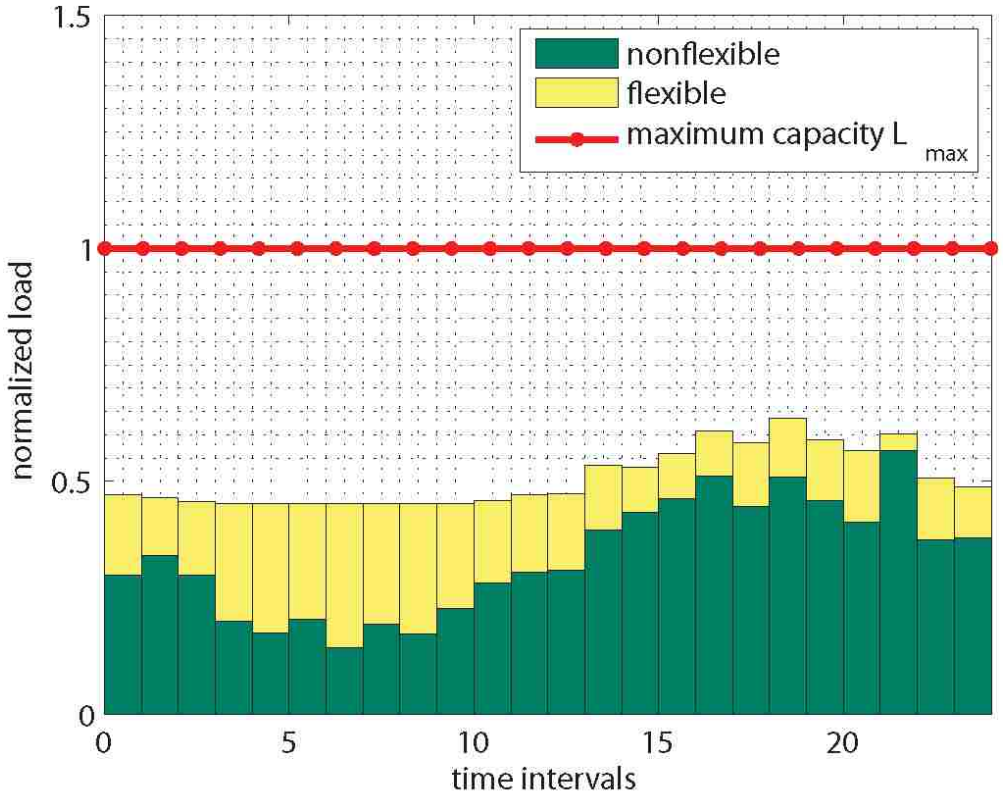

Figure 3.8: The non-flat one-dimension profile as a result of summing the two-dimension water-filling profile over all customers.

$$C^* = \min_{i,k} \sum_{k=1}^{K} \sum_{i=1}^{I} \alpha(l_{i,k}^N + x_{i,k})^2 \tag{3.9}$$

For comparison purpose, we consider another extreme case in which there is only one customer in each cluster. Denote by $\tilde{C}^*$ the minimum cost corresponding to the scenario that every single customers is one cluster, which is defined as

$$\tilde{C}^* = \sum_{k=1}^{K} \tilde{C}_k^* = \sum_{k=1}^{K} \min_{x_{i,k}} \sum_{i=1}^{I} \alpha(l_{i,k}^N + x_{i,k})^2 \tag{3.10}$$

If we assume that the generation capacity of the Utility is higher than the maximum total load requests from all customers at each time instant, the optimization problem (3.10) can be decoupled into K individual optimization problems, each corresponding to one cluster. For each cluster, we solve an optimization problem of the following form

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \tilde{C}_k^* = \min_{x_{i,k}} \sum_{i=1}^{I} \alpha (l_{i,k}^N + x_{i,k})^2 \quad , \\ \text{subject to} \quad & -x_{i,k} \leq 0, \quad i = 1, 2, \ldots, I \quad , \\ & \sum_{i=1}^{I} x_{i,k} - l_F = 0 \quad . \end{aligned} \tag{3.11}$$

These decoupled optimization problems can be solved by using the same technique and the solution structure corresponding to (3.10) is given by a two dimensional water-filling result but with different water levels for different clusters (different customers in this case), as shown in Fig. 3.9.

In general, the minimum cost increases as the size of the cluster increases from 1 to K, so that $C^* \leq \tilde{C}^*$. This performance degradation of the distributed solution comes from the fact that, instead of the global cost, costs of clusters are minimized. We refer to this performance degradation as the Price of Anarchy (PoA). This inefficiency can be characterized by the quantity $\frac{\tilde{C}^* - C^*}{C^*}$, which is the normalized extra cost of opting for distributed objectives over the global objective. Fig. 3.10 shows the normalized extra cost as the customer cluster size increases from 1 to 20. A local water-filling solution applies within each customer cluster. It can be observed that the POA decreases monotonically as the group size increases.

Based on the analysis presented above on the price-based DR scheme we proposed, it will be very interesting to further generalize the problem by assuming that all distributed customers have the decision making capabilities and compete

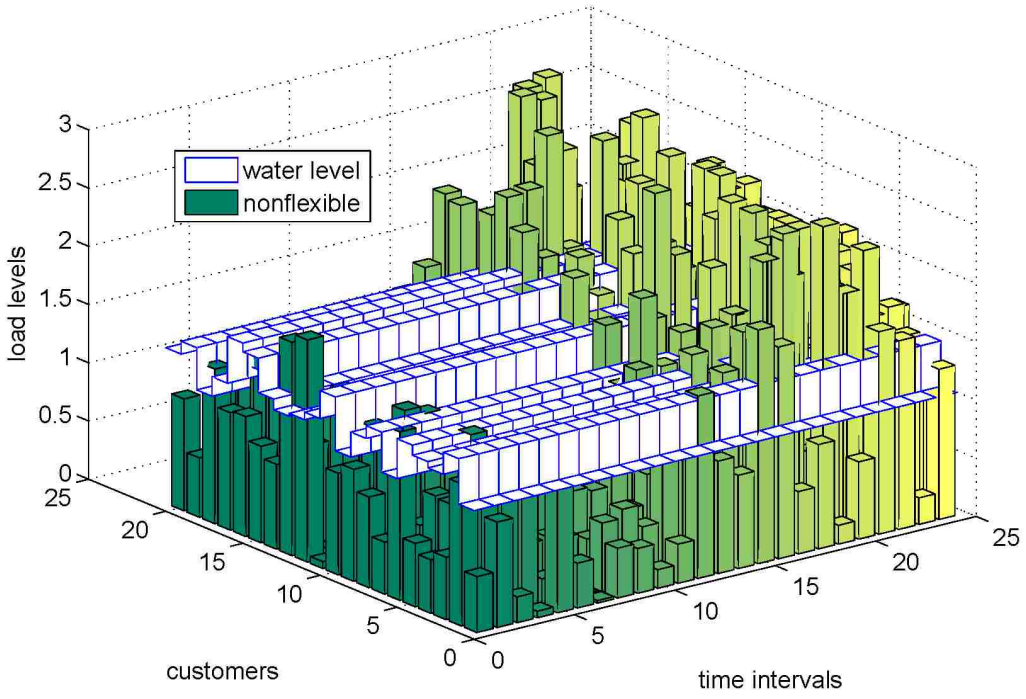

Figure 3.9: Two-dimensional water-filling with different water levels for different customers.

for electric power. In this scenario, the optimization problem becomes a game and a Bayesian Nash equilibrium is needed to achieve the optimal DR scheduling and power allocation. In [31], we presented a Vickrey auctioning game combining the two objective functions of Utility cost minimization and customer profit maximization. Due to space limit, we are going into the details of the game theoretic approach in this paper. Interested readers are referred to [31] for more details.

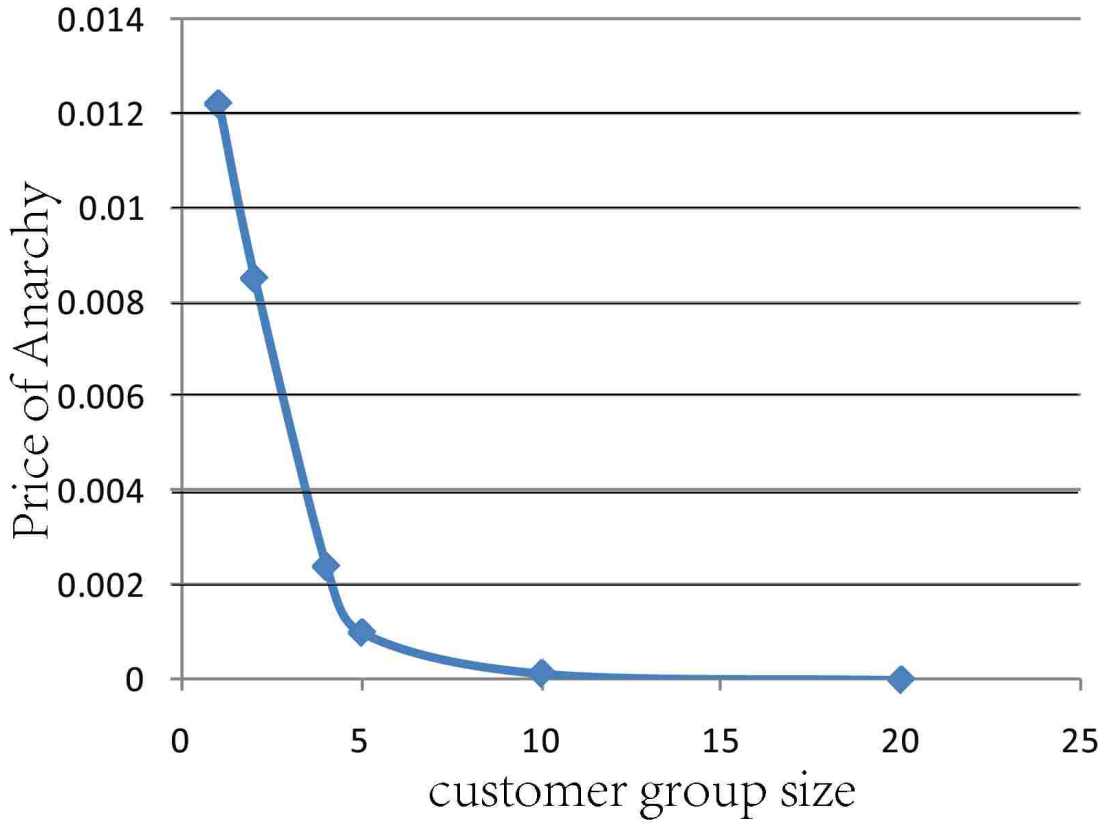

Figure 3.10: Price of anarchy (POA) decreases as the size of the customer cluster increases.

3.6 Conclusion

In this chapter, based on the proposed load demand prediction techniques, we designed a DR scheduling scheme based on the Utility cost minimization with different customer clustering sizes. A convex optimization problem was formulated and the optimal demand response profile was in the form of a two-dimensional *water-filling solution* either with flat water levels or different water levels for different customers. Price of Anarchy (PoA) analysis was presented to balance both the centralized and distributed competing objectives.

Chapter 4

Optimal Stochastic Tracking for Primary Frequency Control

4.1 Introduction

In traditional electric grid planning, the Utility and customers interact on a slow time scale due to insufficient information exchange between the power generation and the consumption sides [2]. Other limiting factors are the uncertainties raised due to the ever-increasing and fluctuating load demands and renewable generations, which are mostly based on solar, wind and tidal resources [19, 67]. Such uncertainties will become more significant as more and more integrated distributed energy resources (DERs), e.g. plug-in hybrid electrical vehicles (PHEVs), are connected to the grid. Thus system operators need more efficient and effective control schemes to balance variables on both generation and consumption sides. These schemes help to overcome many technical challenges with increased penetration of renewable generations and PHEVs, such as voltage rise effects, power quality and power grid protection.

In any electric system, the stability of the electrical grid is guaranteed by bal-

ancing the power generation and consumption [28]. Generation units and even load in some cases must be manipulated to conduct power balancing so the network user is not affected by load changes or generation and transmission outages. From the viewpoint of load matching, various demand response schemes have been proposed to affect customer load profiles [23, 29–31]. In [32], a three-step methodology was presented to manage the cooperation among technologies of distributed generation, distributed electricity storage and demand side load management. From the viewpoint of power generation control, since massive storage of alternating electricity is difficult, two separate equilibria should be kept on the grid for stabilizing purpose [33]: (1) The active power generated should at each moment equal the active power consumed. A deviation from this equilibrium results in a deviation from the standard frequency (60 or 50 Hz). Hence, keeping this equilibrium between active power consumption and generation means maintaining frequency. (2) The reactive power on the grid should be kept in equilibrium as well. Reactive power is an extra load for the grid, leaving less capacity for active power, resulting in a local voltage drop. Hence, keeping reactive power in equilibrium means maintaining voltage. Studies on frequency and voltage control have been reported in many previous work [28, 34–37]. In particular, a comprehensive survey on frequency and voltage control technical features can be found in [28]. In [34], the authors discuss the issue of excess steady-state voltage rise and the methods of limitation that can be applied with specific reference to wind generation. In [36], a strategy for the control of terminal voltage and frequency of a stand-alone self-excited induction generator-(SEIG) based wind generator, working with variable speed and load is proposed. In [37], the authors presented a micro hydro scheme with parallel operation of synchronous and induction generators in micro hydro scheme.

In most of the literature, frequency and voltage control schemes are usually designed separately because generally they are implemented by generator rotor speed governor and excitation control system respectively [8,38]. In this paper, we focus on

frequency control (active power control) issues. The frequency control usually consists several layers [6–8], including primary control, secondary control, tertiary control and other possible balancing power reserve planning services. Control schemes of different levels have different objectives and operating time scales, as shown in Fig. 4.1. The objective of primary frequency control (with a controlling period on the order of seconds) is to maintain a balance between generation and consumption within the synchronous area using turbine speed or turbine governors. However, primary frequency control stabilizes frequency but does not drive the system frequency back to the original set-point value after a disturbance. Secondary frequency control (with a controlling period on the order of minutes) is needed since when several generators are doing generation sharing, secondary frequency control distributes the power imbalance among selected units [6]. The secondary frequency control can also drive the system frequency back to the original desired value. Tertiary frequency control (with a controlling period in the order of minutes to hours) is a manual change in the dispatching in order to restore the secondary reserve and provide a more permanent solution if the imbalance between consumed power and scheduled power persists. There are several important research issues associated with both secondary and tertiary frequency control, such as *spinning reserve*, *unit commitment* and *economic dispatch*. Spinning reserve [42] is the unused capacity provided by devices that are synchronized to the network and can be quickly activated on decision of the system operator. Unit commitment and economic dispatch [43] is to find the optimal dispatch of available generation resources to meet the electrical load and spinning reserves. Other layers in the frequency control framework include stand-by supplies and contractual load shedding which have longer control periods (hours). Unit commitment and economic dispatch are important topics in power grid generation planning by themselves and will not be discussed in detail here.

Currently the most widely adopted primary frequency control scheme is the proportional-integral-derivative controller (PID controller) [6]. This is because PID

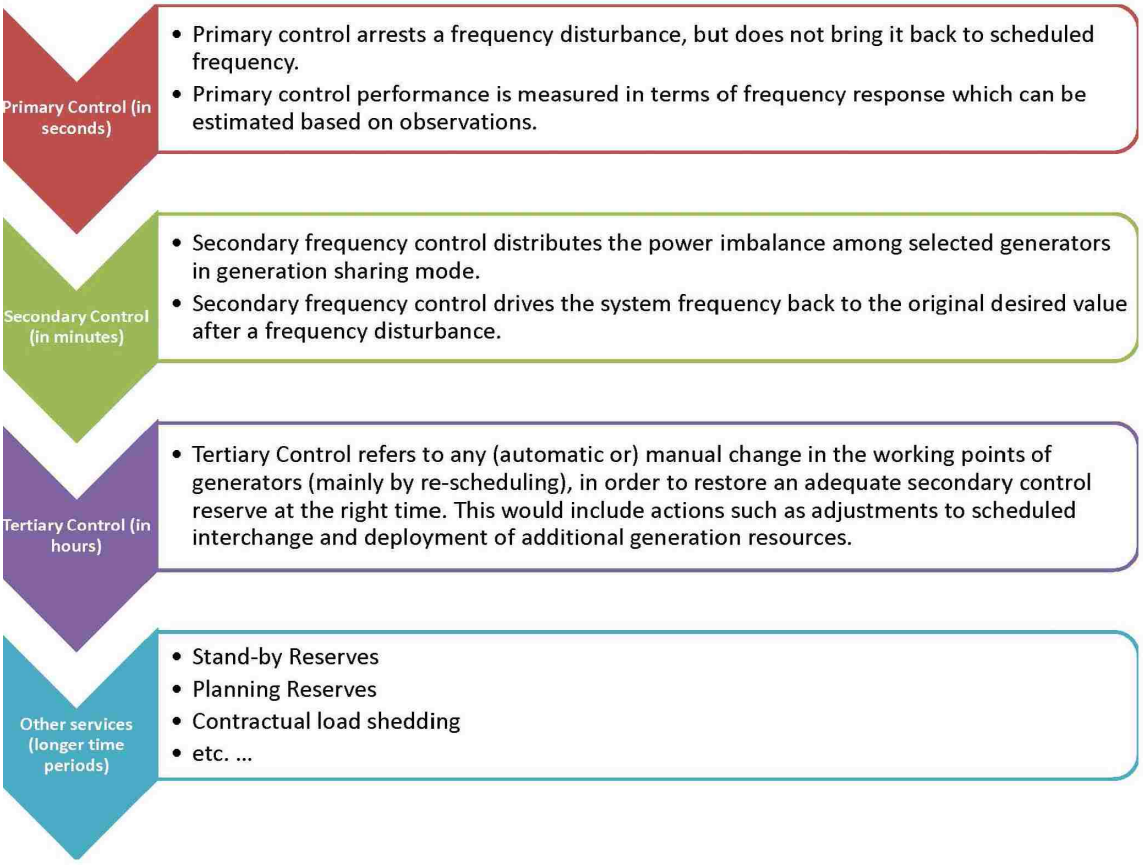

Figure 4.1: Load frequency control consists of primary control, secondary control, tertiary control and other planning reserve services.

controller shows relatively good control performance when the dynamics of the plant is unknown or too complicated to analyze. By tuning the three parameters in the PID controller algorithm, the controller can conveniently provide control action designed for specific process requirements. However, PID controller does not guarantee optimality in control and system stability. In this paper, we focus on the primary frequency control design and propose an optimal stochastic tracking scheme for synchronous generator active power generation control, assuming the dynamics of individual synchronous generator. In this tracking scheme we minimize the difference between the active power generation output and the reference signal which

incorporate the randomness of both load demands and renewable generations. Further analysis on the tracking performance are presented considering synchronous and asynchronous customer load signals in the reference.

The rest of the chapter is organized as follows: In section 4.2, we propose a comprehensive real-time interactive framework for smart-grid while ensuring grid-stability and Quality-of-Service (QoS). In section 4.3 emphasis is placed on developing dynamic models taking into account the uncertainties in renewable distributed generations and customer load demands. In section 4.4, based on the dynamical model of the synchronous generator, two stochastic tracking control schemes are proposed: (1) reference dynamics-based tracking and (2) reference statistics-based tracking. The proposed tracking schemes are further extended by introducing asynchronous customer load demand signals in section 4.5. Simulation results are presented in section 4.6 showing the performances of both prediction and tracking schemes. The conclusions from this chapter are given in section 4.7.

4.2 A Comprehensive Interactive Architecture for Smart Grid

Although a comprehensive formulation and an analysis is not yet available for smart-grid, still there have been several attempts to understand, model and analyze various aspects of smart-grid in [4, 5, 68]. An adaptive stochastic control framework was presented in [69] which was mainly focused on self-healing, prediction and cyber-security of power grids. In [68], the author presented the concept of energy internet, which modeled energy flows from suppliers to customers as data packets in the Internet. The uncertainty in supply due to these integrated renewable DERs and the challenges they imposed on the existing distribution infrastructure and the system

operator have been discussed in [9]. The distribution-level smart-grid features such as interconnection of distributed generation and active distribution management, advanced metering infrastructure (AMI) systems in network management and power quality monitoring were discussed in [10]. In [11], the implementation of vehicle-to-grid (V2G) power issues, strategies and business models for doing so, for purposes of both stabilizing the grid and supporting large-scale renewable energy were discussed. Various control-theoretic and system-level problem formulations of smart-grid architectures have been discussed in [70, 71]. In [12], the authors showed that significant improvements can be made to the operations of a smart-grid by providing information about the likely behavior of renewable energy through both online short-term forecasting and longer term assessments. In [13], a distributed control method was proposed for converter-interfaced renewable generation units with active filtering capability. It is worth pointing out, however, that little work has considered a comprehensive cycle of interactions between the Utility and the smart homes taking into account all aspects of customer-side decision making, Utility-side demand response scheduling, distributed energy resources (DERs) for grid-stability and the effects of ICT infrastructure on these.

In this work, we propose a comprehensive real-time interactive framework for smart-grid while ensuring grid-stability and Quality-of-Service (QoS). This control scheme takes into account the intermittent and random nature of renewable generations and individual customer load demands. Figure 4.2 shows the interaction framework, extending the framework in [72], between distributed customers and the Utility, addressing the demand response scheduling and real-time power generation control respectively.

This comprehensive scheme would give rise to a variety of questions for both the Utility as well as for the individual smart-homes in how to best maximize local goals. The overall framework can be implemented in three steps, as shown in Fig. 4.2.

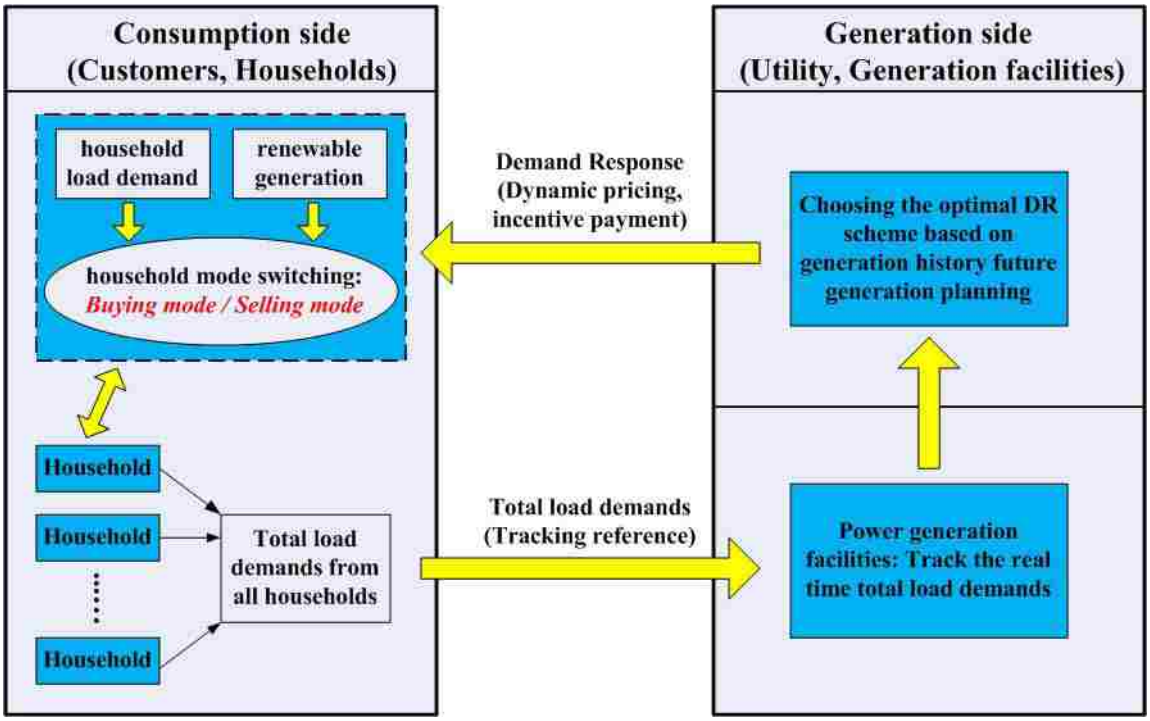

Figure 4.2: Interaction framework for distributed customers and supporting conventional generation facilities.

1. At the beginning of each scheduling interval, the Utility broadcasts electricity price information for several time intervals in advance. Based on the information on prices, as well as predicted load demands, local renewable generation and budgets, customers submit their initial load demand requests to the Utility for each time interval of the scheduling period. Having on-site renewable DERs, each customer is able to at least partly support its own load demand by consuming locally generated renewable energy. When a customer has excess renewable generation, it may decide what portion of its locally generated renewable energy is to be sold to the grid and what portion to be stored in its own distributed storage for future use, in order to optimize long-term accumulated benefit.

2. Based on the initial load demands from customers, the Utility schedules the

Demand Response (DR) for load reallocation over time intervals and among different customers. The objective here is to effectively reshape the load demand profiles by shaving peak loads and flattening the load profiles. In [31], we propose an auctioning game based DR scheme, which integrates minimizing the Utility's cost and maximizing the social welfare of competing customers.

3. In each time interval, the primary control scheme is implemented for each generator to match the active power generation to the reference load assigned for frequency control purpose. For an individual generator, the reference signal for active power generation is random and time-varying. This defines a stochastic reference tracking control problem for synchronous generator in the presence of transmission and sensor measurement feedback delays and errors as well as incomplete knowledge on plant models.

With the first two steps addressed in our previous work [16, 30, 31, 72–74], in this paper we focus on step 3 on developing optimal tracking control schemes for generation facilities to track the time variant load demands. The tracking scheme design becomes challenging when the demand signals from distributed customer locations experience different time delays during data measurement at sensors and data transmission over communication channels, thus becoming asynchronous.

4.3 Dynamical Models for Central Power Plant and Customer Net Load Demands

4.3.1 State-space Representation of a Synchronous Generator

As mentioned in previous sections, the active power imbalance is distributed among selected generators in generation-sharing mode. In primary frequency control stage, we focus on the real-time control scheme design (with controlling periods on the order of seconds) for individual generator, whose dynamics is fast enough to respond to variations in load demands and renewable generations. In this section, we assume the commonly used third-order nonlinear model of a synchronous generator [75]. The nomenclature is defined as follows: r_d, r_q and $r_d^{'}$ are the augmented reactance, i.e. the line and transformer reactance are added with them. δ is the rotor angle with respect to the machine terminal. ω is the relative speed of the rotor in rad/s. $v_q^{'}$ is the transient internal voltage of armature. E_{FD} is the equivalent electromotive force (EMF) in the excitation coil. i_d, i_q are the direct and quadrature axis stator currents. J, I are the rotor inertia and the damping factor. P_e is the terminal active power. $T_{do}^{'}$ is the direct-axis transient time constant. T_e is the output electric torque. T_m is the input mechanical torque. v_t is the generator terminal voltage.

With all parameters above defined in per unit values, the third order nonlinear model is described by the following equations [76]:

$$\dot{\delta} = \omega, \quad \dot{\omega} = \frac{1}{J}\left(T_m - T_e - T_D\right), \quad \dot{v}_q^{'} = \frac{1}{T_{do}^{'}}\left(E_{FD} - v_q^{'} - \left(r_d - r_d^{'}\right) i_d\right) \quad (4.1)$$

where

$$i_d = \frac{v_q^{'} - v_t \cos\delta}{r_d^{'}}, \quad i_q = \frac{v_t \sin\delta}{r_q}, \quad \dot{v}_q^{'} = \frac{1}{T_{do}^{'}}\left(E_{FD} - v_q^{'} - \left(r_d - r_d^{'}\right) i_d\right) \tag{4.2}$$

For a single synchronous generator, it is assumed that the field voltage, rotor angle and the electrical power can be measured. Thus, we define the system state x_c, input u_c and output y_c as $x_c = [x_1, x_2, x_3]^T = [\delta, \omega, v_q^{'}]^T$, $u_c = [u_1, u_2]^T = [E_{FD}, T_m]^T$ and $y_c = P_e$.

By linearizing the nonlinear equations above near a certain operating point "o", we have the linear system and output equations as

$$\dot{x_c} = A_c x_c + B_c u_c, \quad y_c = C_c x_c, \tag{4.3}$$

where

$$A_c = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{K_1}{J} & -\frac{I}{J} & -\frac{K_2}{J} \\ -\frac{K_4}{T_{do}^{'}} & 0 & -\frac{1}{K_3 T_{do}^{'}} \end{bmatrix}, \quad B_c = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{J} \\ \frac{1}{T_{do}^{'}} & 0 \end{bmatrix}, \quad C_c = \begin{bmatrix} K_1 & 0 & K_2 \end{bmatrix} \tag{4.4}$$

The parameters are defined as follows:

$$K_1 = \frac{v_t}{r_d^{'}} x_{3o} \cos x_{1o} + \frac{v_t^2}{2}\left(\frac{1}{r_q} - \frac{1}{r_d^{'}}\right) \sin\left(2x_{1o}\right), \tag{4.5}$$

$$K_2 = \frac{v_t}{r_d^{'}} \sin x_{1o}, \quad K_3 = \frac{r_d^{'}}{r_d}, \quad K_4 = \left(r_d - r_d^{'}\right) \frac{v_t \sin x_{1o}}{r_d^{'}} \tag{4.6}$$

Considering most controllers are implemented digitally, usually with microprocessors, we can further get the discrete time system equations by sampling the original continuous time system. The dynamic equations are given by $x(i+1) = Ax(i) + Bu(i)$ and $y(i) = Cx(i)$, where $x(i) = x_c(kh)$, $u(i) = u_c(kh)$, $y(i) = y_c(kh)$, $A = e^{A_c h}$, $B = \int_0^h e^{A_c \tau} B_c \mathrm{d}\tau$ and $C = C_c$.

4.3.2 Time Varying Autoregressive (TVAR) Process for Customer Load Demand Modeling

Two approaches are widely adopted in the literature for load demand modeling [22, 73, 77–79]. The first approach is component-based load modeling, which reconstructs the expected daily electrical loads of a customer based on appliance sets, occupancy patterns, and statistical data. For example, in [20], the authors constructed such electric load profiles from individual appliance profiles. The second approach is termed the measurement-based load modeling. In [22], a methodology of measurement-based load modeling for transient stability analysis was proposed and Genetic Algorithms (GA) was used to estimate the model parameters. Considering the time correlation in customer power consuming behavior, in this work, we develop a time series analysis based approach for the customer load demand modeling, which provides an efficient way to predict the customer load demand directly based on the immediate load data history.

With the TVAR model for the household load demands in place, we can define the state $x_l(i) = [s(i-1), s(i-2), \ldots, s(i-p)]^T$, so that a state-space representation of the customer load demand can be written as

$$\begin{aligned} &x_l(i+1) = A_l(i)x_l(i) + w_{l1}(i), \\ &z_l(i) = C_l(i)x_l(i), \quad y_l(i) = z_l(i) + w_{l2}(i) \end{aligned} \tag{4.7}$$

The parameters are given by

$$A_l(i) = \begin{bmatrix} \phi_1(i) & \phi_2(i) & \dots & \phi_p(i) \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad w_{l1}(i) = \begin{bmatrix} v(i) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{4.8}$$

$$C_l = [\phi_1(i), \phi_2(i), \dots, \phi_p(i)], \quad w_{l2}(i) = v(i) \tag{4.9}$$

4.3.3 Renewable Generation Modeling

The intermittence of renewable generation mainly comes from the uncertainty of environment, such as variations of wind speed, solar irradiation and cloud movement. Modeling these weather factors are difficult by themselves and are out of the scope of this paper. Thus, here we give a brief overview of related models. Wind speed distributions are often characterized by Weibull or Rayleigh distributions [47]. Historical hourly data for the wind farm site collected over a significant time are normally required to obtain the shaping parameters. In [18], the wind speed probability distributions obtained for the three diverse geographic locations in Canada, are close to Gaussian distributions. Beta distribution validated by different researches as a simple and sufficiently flexible two-parameter distribution, fits well the empirical data of solar irradiation in many situations [19].

The active power generated by conventional generator, modeled by (4.3), needs to track a desired reference signal to keep the system frequency stable. This reference signal is defined as a fraction of the net load demands (the difference between the total load demands and local renewable generations) from local customers being supported by all generators considered. Thus, the dynamics and the statistics of the stochastic reference signal can be obtained from the associated information of customer load demands and renewable generations. Depending on what information

(dynamics or statistics) is indeed available, we may develop different tracking control schemes for primary frequency control in the following section.

4.4 Stochastic Tracking for Primary Frequency Control of the Power Generation Facility

As demonstrated in previous sections, The objective of primary frequency control is to maintain a balance between generation and consumption (demand) and stabilizes the system frequency at a stationary value[1] after a disturbance or incident in the time-frame of seconds, but without restoring the system frequency to the nominal value. Thus the active power generated by the conventional generator is supposed to track the reference signal which depends on the total net load demands of local customers. The tracking control problem is shown in Fig. 4.3, where z_r is the reference signal for active power generation. It is worth pointing out that in a general multi-layer frequency control framework, the deviation of terminal frequency from the frequency set point (50 or 60 Hz) is usually used as the feedback signal in the control loop for both primary and secondary control. This is definitely necessary for secondary control because its most important objective is to restore the system frequency from a temporarily stable level resulting from the primary control back to the nominal frequency level. In traditional primary frequency control, the frequency deviation is also used to define the droop which is the controller gain in the feedback loop [28]. This droop is a function of both the deviation in frequency and the nominal generator output power (reference signal of active power generation). However adjusting the droop settings is not always easy because it often requires that the

[1]Generally speaking, the steady-state frequency deviation level depends on both the primary frequency control of the generation unit and the frequency sensitivity of the load. Since these two factors are decoupled, we ignore the frequency sensitivity of the load and focus on the control design of the primary frequency control.

plant be shut down. Hence, for the primary control stage only, we directly consider the deviation between active power generation and consumption as the feedback signal, which is possible based on the active power output measurement [80, 81]. By doing this we incorporate the uncertainties of load demands and renewable generations in the active power generation reference signal and balance the power without relying on system tests operated by Transmission System Operator (TSO). In this section, we consider the scenario in which the load demand signals from distributed customers are fully synchronous, meaning the net load demand signals from different customers experience the same time delay. Perfect knowledge of the reference signal is also assumed. We will consider different optimal controller design for asynchronous scenario in the next section.

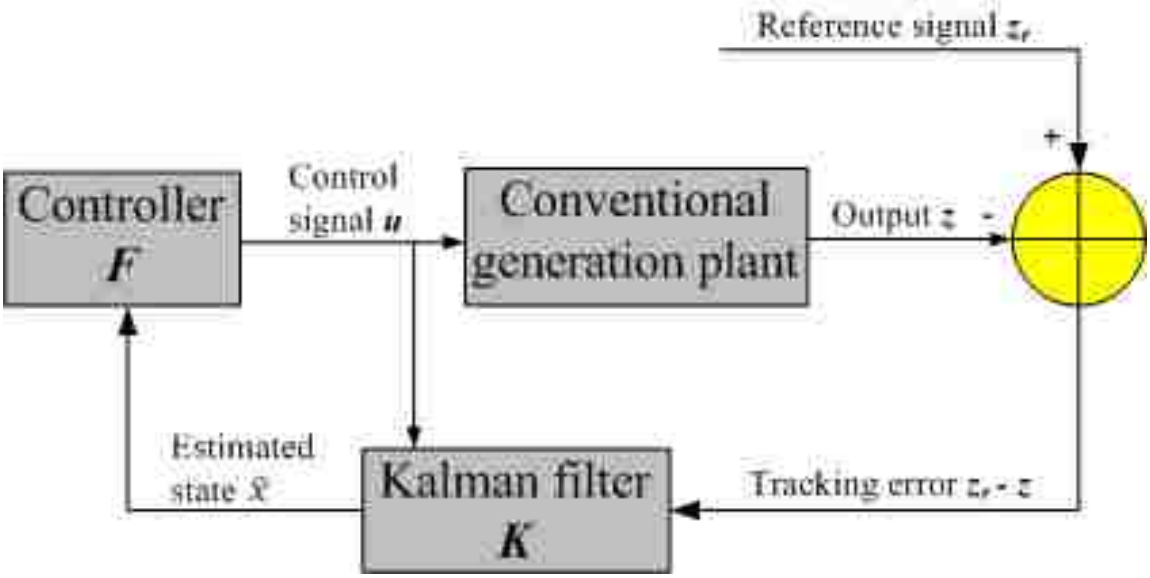

Figure 4.3: Tracking control diagram: The power generation control is implemented by feedback control design. z is the system output, which is the active power generation. The deviation in active power and load serves as the feedback signal. A Kalman filter is adopted to estimate the system state based on the noisy and incomplete observation of the system output.

4.4.1 Reference-Dynamics based Tracking

Recall that the state-space representation of the dynamical system of the conventional synchronous generator can be written as

$$\begin{aligned} x(i+1) &= Ax(i) + Bu(i) + w_1(i), \\ z(i) &= Cx(i), \quad y(i) = z(i) + w_2(i), \end{aligned} \tag{4.10}$$

where $z(i)$ is the system output and $y(i)$ is the noisy output measurement when we try to access the active power output. Process noise w_1 and measurement noise w_2 are both zero mean white noises with auto-covariance matrix intensities V_1, V_2. The cross covariance matrix of w_1 and w_2 is V_{12}. The initial state $x(0)$ is a random vector with mean $\bar{x}_0$ and covariance matrix Q_0. The control interval is $i \in [i_0, i_1]$.

To be specific, and to keep the discussion within the available space, let us assume that the renewable DERs of interest are wind turbines at distributed customer locations. Other renewable sources with dynamical generation, such as tidal resources, can be incorporated in a similar way. Since synchronous generators can also be used in a wind-energy plant, the state-space representation can still be used. Due to the geographic diversity, wind turbines at different locations may have various generator inputs. If we assume that within the controlling period the wind changes rapidly (which leads to uncorrelated increments) and is omni-directional (justifying zero mean), then the input of each wind turbine can be approximated as being white Gaussian [1]. Thus a linear dynamical system for the renewable generation can be written as

[1]Gaussian distribution has been widely used in the literature for wind speed modeling [17,18]. For very short time scale, the assumption of uncorrelated increments in wind speed is relatively strong. But as the time scale of controlling period increases, this assumption becomes more and more reasonable.

$$
\begin{aligned}
&x_w(i+1) = A_w x_w(i) + w_{w1}(i),\\
&z_w(i) = C_w x_w(i), \quad y_w(i) = z_w(i) + w_{w2}(i).
\end{aligned} \tag{4.11}
$$

With the dynamic models of renewable generator (4.11) and customer load demand (4.7) developed above, the net load demand, denoted by y_{net} and $y_{net} = y_l - y_w$, of a single customer can be written as the output of the following dynamic system

$$
\begin{aligned}
\underbrace{\begin{bmatrix} x_l(i+1) \\ x_w(i+1) \end{bmatrix}}_{x_{net}(i+1)} &= \underbrace{\begin{bmatrix} A_l & \mathbf{0} \\ \mathbf{0} & A_w \end{bmatrix}}_{A_{net}} \underbrace{\begin{bmatrix} x_l(i) \\ x_w(i) \end{bmatrix}}_{x_{net}(i)} + \underbrace{\begin{bmatrix} w_{l1}(i) \\ w_{w1}(i) \end{bmatrix}}_{w_{net1}(i)},\\
z_{net}(i) &= z_l(i) - z_w(i) = \underbrace{[C_l, -C_w]}_{C_{net}} \begin{bmatrix} x_l(i) \\ x_w(i) \end{bmatrix},\\
y_{net}(i) &= z_{net}(i) + \underbrace{w_{l2}(i) + w_{w2}(i)}_{w_{net2}(i)}.
\end{aligned} \tag{4.12}
$$

The reference signal of individual synchronous generator, which depends on the total net load demand (reference signal for individual generator is assigned by TSO), can be written as the output of a dynamical model defined in (4.13), by doing system augmentation over independent customers.

$$
\begin{aligned}
&x_r(i+1) = A_r x_r(i) + w_{r1}(i),\\
&z_r(i) = C_r x_r(i), \quad y_r(i) = z_r(i) + w_{r2}(i).
\end{aligned} \tag{4.13}
$$

All the other quantities V_{r1}, V_{r2} and V_{r12} are similarly defined with the ones for synchronous generators. Thus, the original frequency regulation problem can

be formulated as an optimal stochastic tracking problem with incomplete and noisy observations, in which we want to minimize the quadratic control cost

$$U = \mathbb{E}\left\{\sum_{i=i_0}^{i_1} \left(z(i) - z_r(i)\right)^T R_1(i) \left(z(i) - z_r(i)\right) + u^T(i) R_2(i) u(i)\right\} \tag{4.14}$$

where the control accuracy measure matrix $R_1(i)$ is positive semi-definite and the control effort measure matrix $R_2(i)$ is positive definite. To solve this problem, we can consider the augmented system:

$$\begin{aligned}
\underbrace{\begin{bmatrix} x(i+1) \\ x_r(i+1) \end{bmatrix}}_{\tilde{x}(i+1)} &= \underbrace{\begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & A_r \end{bmatrix}}_{\tilde{A}} \underbrace{\begin{bmatrix} x(i) \\ x_r(i) \end{bmatrix}}_{\tilde{x}(i)} + \underbrace{\begin{bmatrix} B \\ \mathbf{0} \end{bmatrix}}_{\tilde{B}} u[i] + \underbrace{\begin{bmatrix} w_1(i) \\ w_{r1}(i) \end{bmatrix}}_{\tilde{w}_1(i)}, \\
\tilde{z}(i) = z(i) - z_r(i) &= \underbrace{[C, -C_r]}_{D} \begin{bmatrix} x(i) \\ x_r(i) \end{bmatrix}, \\
\underbrace{\begin{bmatrix} y(i) \\ y_r(i) \end{bmatrix}}_{\tilde{y}(i)} &= \underbrace{\begin{bmatrix} C & \mathbf{0} \\ \mathbf{0} & C_r \end{bmatrix}}_{\tilde{C}} \underbrace{\begin{bmatrix} x(i) \\ x_r(i) \end{bmatrix}}_{\tilde{x}(i)} + \underbrace{\begin{bmatrix} w_2(i) \\ w_{r2}(i) \end{bmatrix}}_{\tilde{w}_2(i)}.
\end{aligned} \tag{4.15}$$

With the augmented system above, the original objective (A.7) can be rewritten as $U = \mathbb{E}\left\{\tilde{z}^T(i) R_1(i) \tilde{z}(i) + u^T(i) R_2(i) u(i)\right\}$, where system state is incomplete, noisy and need to be reconstructed based on the measurements. The original tracking problem has been converted into an output regulator problem of the augmented system (4.15), in which $\tilde{z}(i)$ is the controlled variable and $\tilde{y}(i)$ is the noisy observation. The initial state and covariance matrix are given by $\tilde{x}_0$ and $\tilde{Q}_0$. We assume that the initial state is uncorrelated with both process noise and measurement noise. Thus, $[\tilde{w}_1(i), \tilde{w}_2(i)]$ is a joint white noise vector process. The covariance matrix intensities of the initial state and white noises are given by

According to the separation principle [82], the solution to this problem is given by $u(i) = -F(i)\hat{\tilde{x}}(i)$, $i = i_0, i_0+1, \ldots, i_1$, where feedback gain is given by

$$\begin{aligned} F(i) = &\left\{R_2(i+1) + \tilde{B}^T[D^T R_1(i+1)D + P(i+1)]\tilde{B}\right\}^{-1} \\ &\cdot \tilde{B}^T\left[D^T R_1(i+1)D + P(i+1)\right]\tilde{A}. \end{aligned} \tag{4.16}$$

The sequence of matrices $P(i)$ satisfy the matrix difference equation

$$P(i) = \tilde{A}^T[D^T R_1(i+1)D + P(i+1)][\tilde{A} - \tilde{B}F(i)], \tag{4.17}$$

with terminal cost weight matrix $P(i_1) = P_1 = D^T R_1(i_1)D$. Furthermore, $\hat{\tilde{x}}(i)$ is the minimum mean square linear estimation of $\tilde{x}(i)$ given $\tilde{y}(i)$, $i = i_0, i_0+1, \ldots, i_1 - 1$. Since we assume the nonsingular case ($\tilde{V}_2(i) > 0$), $\hat{\tilde{x}}(i)$ can be obtained as the output of the optimal observer [82]:

$$\hat{\tilde{x}}(i+1) = \tilde{A}\hat{\tilde{x}}(i) + \tilde{B}u(i) + K(i)\left[\tilde{y}(i) - \tilde{C}\hat{\tilde{x}}(i)\right]. \tag{4.18}$$

Denoting by $e(i)$ the reconstruction error $e(i) = \tilde{x} - \hat{\tilde{x}}$, then the optimal observer minimizes the mean square reconstruction error $\mathbb{E}\{e^T(i)W(i)e(i)\} = tr\,[Q(i)W(i)]$ for any predefined positive definite matrices $W(i)$. The optimal gain matrices $K(i)$ can be obtained from the recurrence relations:

$$K(i) = \left[\tilde{A}Q(i)\tilde{C}^T + \tilde{V}_{12}(i)\right]\left[\tilde{V}_2(i) + \tilde{C}Q(i)\tilde{C}^T\right]^{-1}, \tag{4.19}$$

$$Q(i+1) = \left[\tilde{A} - K(i)\tilde{C}\right]Q(i)\tilde{A}^T + \tilde{V}_1(i) - K(i)\tilde{V}_{12}^T(i), \tag{4.20}$$

where $Q = \mathbb{E}\left\{(\tilde{x} - \hat{\tilde{x}})(\tilde{x} - \hat{\tilde{x}})^T\right\}$ is the second order moment matrix of the reconstruction error with the initial value $Q(i_0) = \tilde{Q}_0$. The initial condition of the observer state is $\hat{\tilde{x}} = \bar{\tilde{x}}_0$.

Figure 4.4 shows the tracking performance averaged over 100 realizations of the stochastic reference. It can be seen that the tracking error decreases to zero over time. In this simulation, the synchronous generator parameters in wind power plant and conventional plant are the same. In practice, matrix parameters in (4.10) can be calculated according to real impedance and admittance values. The parameters in the dynamic system of load demands are the same with the TVAR parameters we obtained in previous sections.

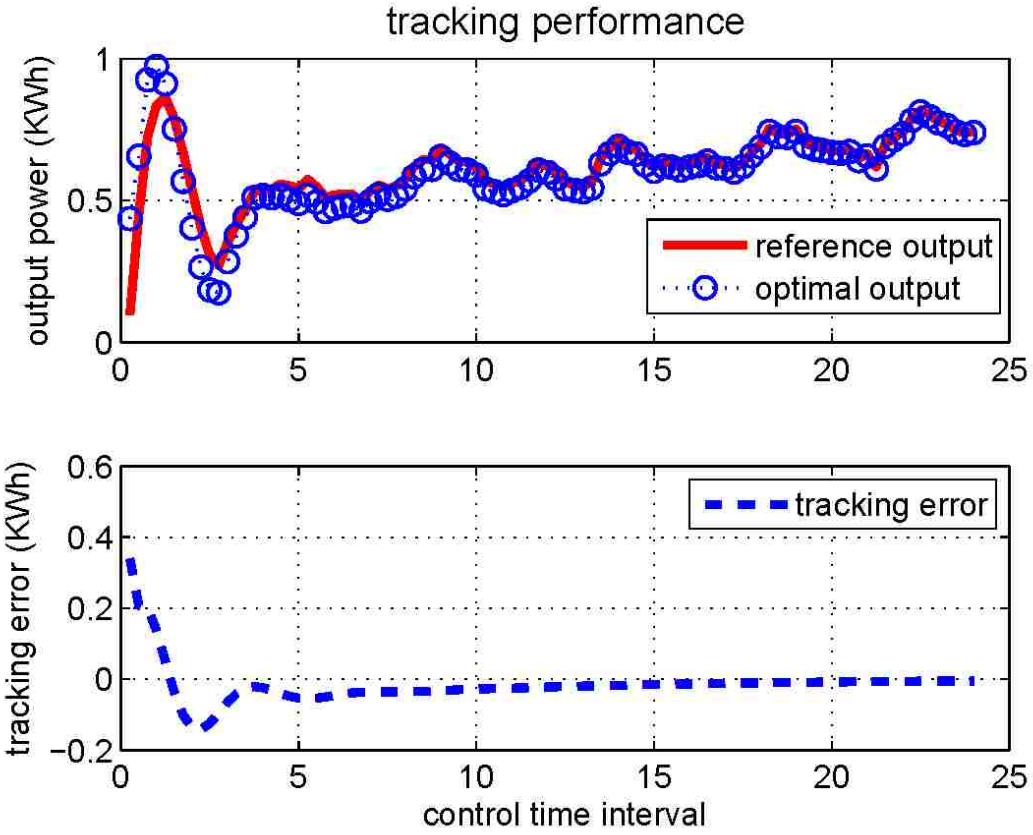

Figure 4.4: Tracking performance averaged over 100 realizations of the stochastic reference. It can be seen that the tracking error decreases to zero over time. In this simulation, the synchronous generator parameters in wind power plant and conventional plant are the same. In practice, parameters can be calculated according to real impedance and admittance values.

4.4.2 Reference-Statistics based Tracking

Note that the tracking scheme proposed above is based on the reference dynamics, which are usually based on the internal structure of renewable generators. In practice, modeling renewable generation as a dynamical process could be difficult. Some renewable generation does not follow a dynamical process by nature. For example, the energy conversion on solar panels from irradiation to electricity falls into the photoelectric effect. On the other hand, the statistics of the distributions of renewable generations and load demands are generally available as we discussed in previous sections. For example, if the renewable generation is predicted from weather forecast data, then the reference signal distributions may be calculated based on customer load distributions obtained from real measured data [83]. To incorporate more renewable resources into the discussion, as well as making the tracking easier to implement in practice, we propose a stochastic based tracking controller based on the mean process and the second order moment of the reference signal, which are assumed to be known at the beginning of the control interval. Denote by $\bar{z}_r$ and M_z the mean process and the second order moment matrix respectively, the control performance index (A.7) can be rewritten as

$$\begin{aligned} U &= \mathbb{E}\left\{\sum_{i=i_0}^{i_1}\left(z(i)-z_r(i)\right)^T R_1(i)\left(z(i)-z_r(i)\right)+u^T(i)R_2(i)u(i)\right\} \\ &= \sum\left\{\mathrm{tr}\left(R_1M_z\right)-\bar{z}_r^TR_1\bar{z}_r+\left(z-\bar{z}_r\right)^TR_1\left(z-\bar{z}_r\right)+u^TR_2u\right\} \end{aligned} \tag{4.21}$$

It can be observed that the optimal control law of this mean square optimization problem is the same to a deterministic tracking problem where the known deterministic state reference signal is the mean process of the stochastic reference. The corresponding deterministic tracking objective can be written as

$$U' = \sum [(z - \bar{z}_r)^T R_1 (z - \bar{z}_r) + u^T R_2 u], \tag{4.22}$$

and we know from [84] that the optimal control law is given by

$$u(i) = -H(i)\hat{x}(i) + H_g(i)g(i+1), \tag{4.23}$$

where

$$H(i) = \left[R_2(i) + B^T P(i+1)B\right]^{-1} B^T P(i+1)A, \tag{4.24}$$

$$H_g(i) = \left[R_2 + B^T P(i+1)B\right]^{-1} B^T. \tag{4.25}$$

The sequence of matrices $P(i)$ satisfy matrix difference Riccati equation with terminal condition $P(i_1) = C^T R_1(i_1)C$:

$$P(i) = A^T P(i+1)\left[I + BR_2^{-1}(i)B^T P(i+1)\right]^{-1} A + C^T R_1(i)C \tag{4.26}$$

$g(i)$ can be obtained by solving the following difference equation with boundary condition $g(i_1) = C^T R_1(i_1)\bar{z}_r(i_1)$:

$$\begin{aligned} g(i) = A^T &\left[I - [P^{-1}(i)(i+1) + BR_2^{-1}(i)B^T]^{-1}\right. \\ &\left. \cdot BR_2^{-1}(i)B^T\right] g(i+1) + C^T R_1(i)\bar{z}_r(i). \end{aligned} \tag{4.27}$$

Based on the separation principle of controller and observer design, the observer can be implemented separately with the form of $\hat{x}(i+1) = A\hat{x}(i) + B\hat{u}(i) +$

$K(i)\,(y(i) - C\hat{x}(i))$. In the presence of observation noise, the optimal observer design is the same as in the case of reference dynamics-based tracking scheme.

The top figure in Fig. 4.5 shows the tracking performance for one realization of the stochastic reference signal. The bottom figure shows the tracking error averaged over 100 realizations. Comparing Fig. 4.4 and Fig. 4.5, we can see the difference in tracking performances. In reference-dynamics based tracking scheme, the tracking error decreases to zero as the tracking proceeds over time. While in reference-statistics based tracking scheme, the tracking error fluctuates around zero after a certain point with a stable average error of around 6%. This is because in reference-statistics based tracking scheme, what is really being tracked is the mean process of the stochastic reference signal. The tracking error will not converge to zero because there is always a deviation between individual realization and the mean process. It is worth pointing out that this is the price we pay for replacing the precise dynamic information with relatively simple statistical information of the reference signal.

4.5 Reference Prediction In the Presence of Asynchronous Load Demand Signals

Note that the two tracking schemes developed above are based on the strong assumption that the perfect reference signal is available for active power control. This means the load demand signals from distributed customers are assumed to be fully synchronous (experiencing the same measurement and transmission delays), which, however, may not hold in practice. In this section, we relax the assumption and investigate the scenario with asynchronous customer load demand signals that experience different time delays arising from sensor measurement and data transmission, as shown in Fig. 4.6. We assume a simple model of an electricity market made of

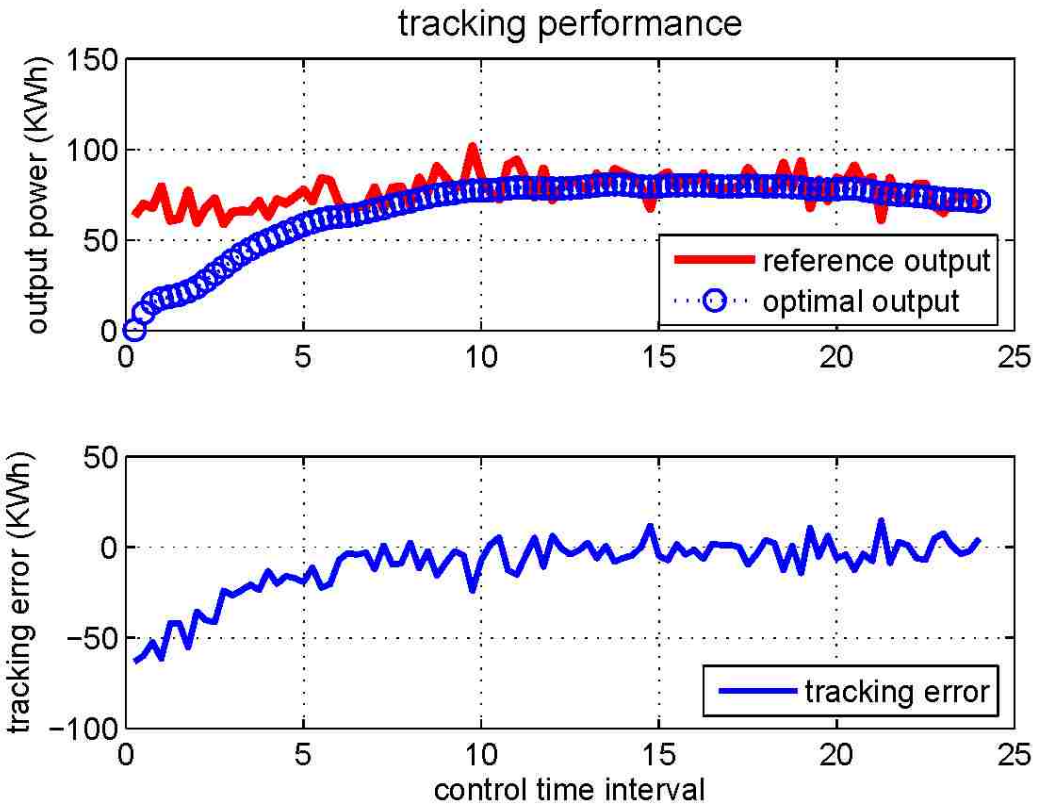

Figure 4.5: Tracking performance based on reference-statistics information. The top figure shows the tracking performance for one realization of the stochastic reference signal. The bottom figure shows the tracking error averaged over 100 realizations.

M customers with renewable DER capabilities integrated in to a grid supported by conventional generation facilities. It is worth pointing out that this simple model may be considered as a building block in a realistic Grid since the real power grid with more complicated and possibly hierarchical structure can be decomposed into smaller units with this type of simple structure. An example is a microgrid with local conventional generation facilities [85]. We can define the dynamical model for customer j, $j = 1, 2, \ldots, M$ in a similar way as we define (4.10) and (4.13).

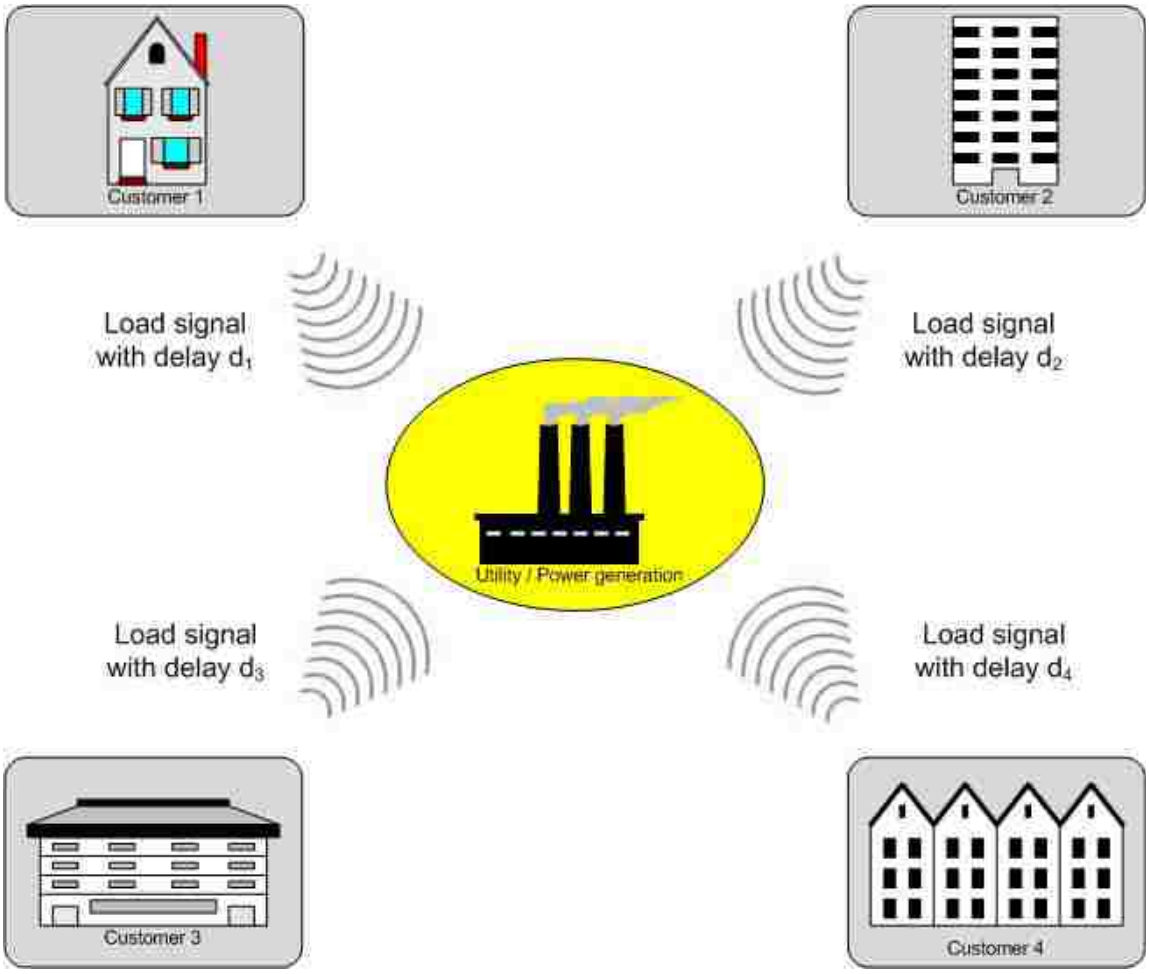

Figure 4.6: Net load demand signals from different customer locations experience different time delays arising from sensor measurement and data transmission.

$$x_{r_j}(i+1) = A_{r_j}x_{r_j}(i) + w_{r_j}(i),$$
$$y_{r_j}(i) = C_{r_j}x_{r_j}(i) + v_{r_j}(i), z_{r_j}(i) = C_{r_j}x_{r_j}(i), \tag{4.28}$$

where $y_{r_j}(i)$ is the net-load demand of the j-th household in time interval i. $w_{r_j}(i)$ and $v_{r_j}(i)$ are process and measurement noises, both of which are assumed to be white Gaussian with covariance matrices Q_{r_j} and R_{r_j}. We assume that $w_{r_j}(i)$, $v_{r_j}(i)$

and the initial state $x_{r_j}(i_0)$ are mutually independent. Thus in the objective function (A.7), we have $z_r(i) = \sum_{j=1}^{M} z_{r_j}(i)$.

Because of the measurement noises, as well as possibly different delays expeXrienced by sensor readings from distributed customers, the load demand signals received by the generation side are asynchronous and noisy versions, denoted by $y_{r_j}(i - d_j)$, where d_j is the delay corresponding to the j-th customer. These asynchronous load demand signals can not be directly added up since they represent load demands in different time intervals. Thus, to solve the original stochastic tracking control problem defined in (A.7), a prediction scheme is required to recover the non-delayed reference signal before any control efforts can be implemented. In this section, we will focus on the reference prediction scheme design, which gives the minimum mean square error (MMSE), to recover the correct reference signal for the power generation tracking control. We start by considering a simplified scenario in which only one customer j is sending the load demand signal, which serves as the reference signal to be tracked by the conventional generator.

4.5.1 Reference Prediction with a Single Customer Load Signal

Based on the customer dynamics (4.28), we define a new dynamical system whose state and output are defined as $x'_{r_j}(i) = [(x_{r_j}(i-d_j))^T, (x_{r_j}(i-d_j+1))^T, \ldots, (x_{r_j}(i))^T]^T$ and $y'_{r_j}(i) = y_{r_j}(i - d_j)$. The system equations are then given by

$$\begin{aligned} x'_{r_j}(i+1) &= A'_{r_j} x'_{r_j}(i) + w'_{r_j}(i), \\ y'_{r_j}(i) &= C'_{r_j} x'_{r_j}(i) + D'_{r_j} v'_{r_j}(i), \\ z'_{r_j}(i) &= z_{r_j}(i) = C_{r_j} x_{r_j}(i) = C''_{r_j} x'_{r_j}(i), \end{aligned} \tag{4.29}$$

where $y'_{r_j}(i)$ is the measurement of the delayed customer demand signal. $z'_{r_j}(i)$ is the non-delayed demand signal and thus is what we want to reconstruct. The system and measurement matrices, as well as the process and measurement noises are given by

$$A'_{r_j} = \begin{bmatrix} A_{r_j} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & A_{r_j} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & A_{r_j} \end{bmatrix}, \quad C'_{r_j} = [C_{r_j}, \mathbf{0}, \dots, \mathbf{0}] \tag{4.30}$$

$$D'_{r_j} = [I, \mathbf{0}, \dots, \mathbf{0}], \quad C''_{r_j} = [\mathbf{0}, \dots, \mathbf{0}, C_{r_j}] \tag{4.31}$$

$$w'_{r_j}(i) = \begin{bmatrix} w_{r_j}(i-d_j) \\ w_{r_j}(i-d_j+1) \\ \vdots \\ w_{r_j}(i) \end{bmatrix}, \quad v'_{r_j}(i) = \begin{bmatrix} v_{r_j}(i-d_j) \\ v_{r_j}(i-d_j+1) \\ \vdots \\ v_{r_j}(i) \end{bmatrix} \tag{4.32}$$

With the system defined above, we can express the estimate of the non-delayed customer signal as a function of the estimate of the state of the new system, given by $\hat{z}'_{r_j}(i) = C''_{r_j}\hat{x}'_{r_j}(i)$. Thus, the original prediction problem has been transformed into an optimal estimation problem of system (4.29).

4.5.2 Kalman Filter Design for Reference Prediction

For system (4.29), the Kalman filter can be used to estimate the state based on the output measurements. Define $Q'_{r_j} = \text{diag}[Q_{r_j}(t-d_j), Q_{r_j}(t-d_j+1), \dots, Q_{r_j}(t)]$ and $R'_{r_j} = R_{r_j}(i-d_j)$ as the covariance matrices of $w_{r_j}(i)$ and $D'_{r_j}v'_{r_j}(i)$ respectively.

Denote by $\hat{x}^{'}_{r_j}(i|i-1)$ the *a priori* state estimate at time i given observations before time i, and by $\hat{x}^{'}_{r_j}(i|i)$ the *a posteriori* state estimate at time i given observations up to and including at time i. These estimates are then given recursively by [86]

$$\hat{x}^{'}_{r_j}(i|i) = \hat{x}^{'}_{r_j}(i|i-1) + K_{r_j}(i)(y^{'}_{r_j}(i) - C^{'}_{r_j}\hat{x}^{'}_{r_j}(i|i-1)) \tag{4.33}$$

$$\hat{x}^{'}_{r_j}(i+1|i) = A^{'}_{r_j}(i)\hat{x}^{'}_{r_j}(i|i) \tag{4.34}$$

where $K_{r_j}(i)$, $i = i_0, i_0+1, \ldots, i_1$ is the Kalman gain matrix. The initialization for the recursion is known and given by $\hat{x}^{'}_{r_j}(0|-1) = E[x^{'}_{r_j}(0)]$. Define the covariance matrix of the prediction error and filtering error as

$$\begin{aligned} P_{r_j}(i|i-1) &= \mathrm{Cov}\{x^{'}_{r_j}(i) - \hat{x}^{'}_{r_j}(i|i-1)\} \\ P_{r_j}(i|i) &= \mathrm{Cov}\{x^{'}_{r_j}(i) - \hat{x}^{'}_{r_j}(i|i)\} \end{aligned} \tag{4.35}$$

The prediction error covariance matrices can be computed jointly with the filtering error covariance matrix from the following recursion with the initialization $P_{r_j}(0|-1) = \mathrm{Cov}\{x^{'}_{r_j}(0)\}$.

$$P_{r_j}(i|i) = P_{r_j}(i|i-1) - K_{r_j}(i)C^{'}_{r_j}P_{r_j}(i|i-1) \tag{4.36}$$

$$P_{r_j}(i+1|i) = A^{'}_{r_j}P_{r_j}(i|i)(A^{'}_{r_j})^T + Q^{'}_{r_j} \tag{4.37}$$

The Kalman gain matrix $K_{r_j}(i)$ is given by

$$K_{r_j}(i) = P_{r_j}(i|i-1)(C^{'}_{r_j}(i))^T \cdot [C^{'}_{r_j}(i)P_{r_j}(i|i-1)(C^{'}_{r_j}(i))^T + R_{r_j}]^{-1} \tag{4.38}$$

Now with the MMSE state estimate $\hat{x}^{'}_{r_j}(i)$ available, we can reconstruct the optimal prediction of the non-delayed reference signal as $\hat{z}_{r_j}(i) = \hat{z}^{'}_{r_j}(i) = C^{''}_{r_j} \cdot \hat{x}^{'}_{r_j}(i)$.

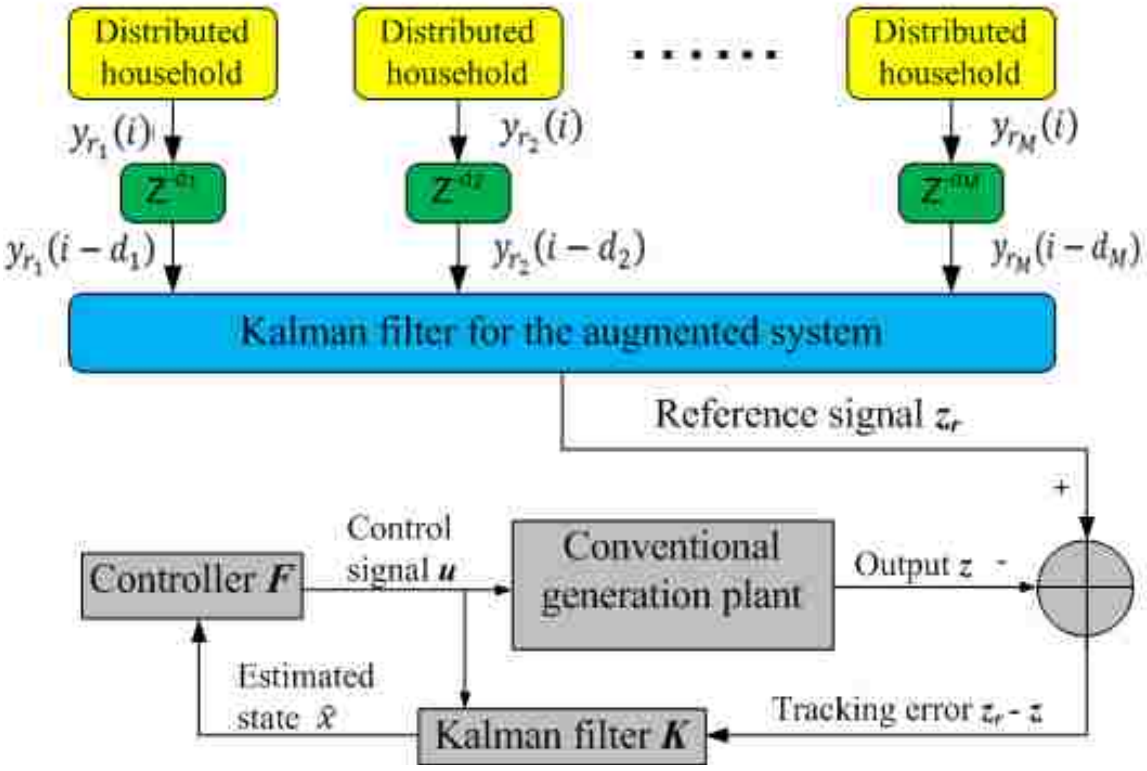

Figure 4.7: The tracking control framework with multiple asynchronous signals from different customers. A centralized Kalman filter is implemented to optimally reconstruct the reference which the conventional active power generation needs to track.

4.5.3 Reference Prediction with Multiple Delayed Household Signals

Based on the results above, we can generalize the reference prediction problem by incorporating multiple asynchronous distributed customer demand signals with different delays. Similarly, we can define the transformed system for each household as in (4.29) and construct an augmented system as follows, with the augmented state and output defined as $x_r(i) = [(x'_{r_1}(i))^T, (x'_{r_2}(i))^T, \ldots, (x'_{r_M}(i))^T]^T$ and $y_r(i) = [(y'_{r_1}(i))^T, (y'_{r_2}(i))^T, \ldots, (y'_{r_M}(i))^T]^T$,

$$x_r(i+1) = A_r x_r(i) + w_r(i), \tag{4.39}$$

$$y_r(i) = C_r x_r(i) + v_r(i), z_r(i) = C_r^{''} x_r(i). \tag{4.40}$$

The system and measurement matrices, as well as the process and measurement noises are given by

$$A_r = \begin{bmatrix} A_{r_1}^{'} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & A_{r_2}^{'} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & A_{r_M}^{'} \end{bmatrix}, \quad w_r(i) = \begin{bmatrix} w_{r_1}^{'}(i) \\ w_{r_2}^{'}(i) \\ \vdots \\ w_{r_M}^{'}(i) \end{bmatrix} \tag{4.41}$$

$$C_r = \begin{bmatrix} C_{r_1}^{'} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & C_{r_2}^{'} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & C_{r_M}^{'} \end{bmatrix}, \quad v_r(i) = \begin{bmatrix} D_{r_1}^{'} v_{r_1}^{'}(i) \\ D_{r_2}^{'} v_{r_2}^{'}(i) \\ \vdots \\ D_{r_M}^{'} v_{r_M}^{'}(i) \end{bmatrix}, \tag{4.42}$$

$$C_r^{''} = [C_{r_1}^{''}, C_{r_2}^{''}, \dots, C_{r_M}^{''}] \tag{4.43}$$

With the linear augmented system well defined, the KF technique can be similarly applied to reconstruct the MMSE estimate of the reference signal as we did for the single customer scenario, which is given by $\hat{z}_r(i) = C_{r_j}^{''} \hat{x}_r(i)$. Fig. 4.7 shows the overall control framework with the prediction block involved. Because of the linearity and the fact that the dynamical systems of different customers are independent with each other, this centralized Kalman filter implementation is equivalent to a distributed scheme in which a distributed KF is implemented for each household demand signal. The reference signal for the central power plant generation tracking

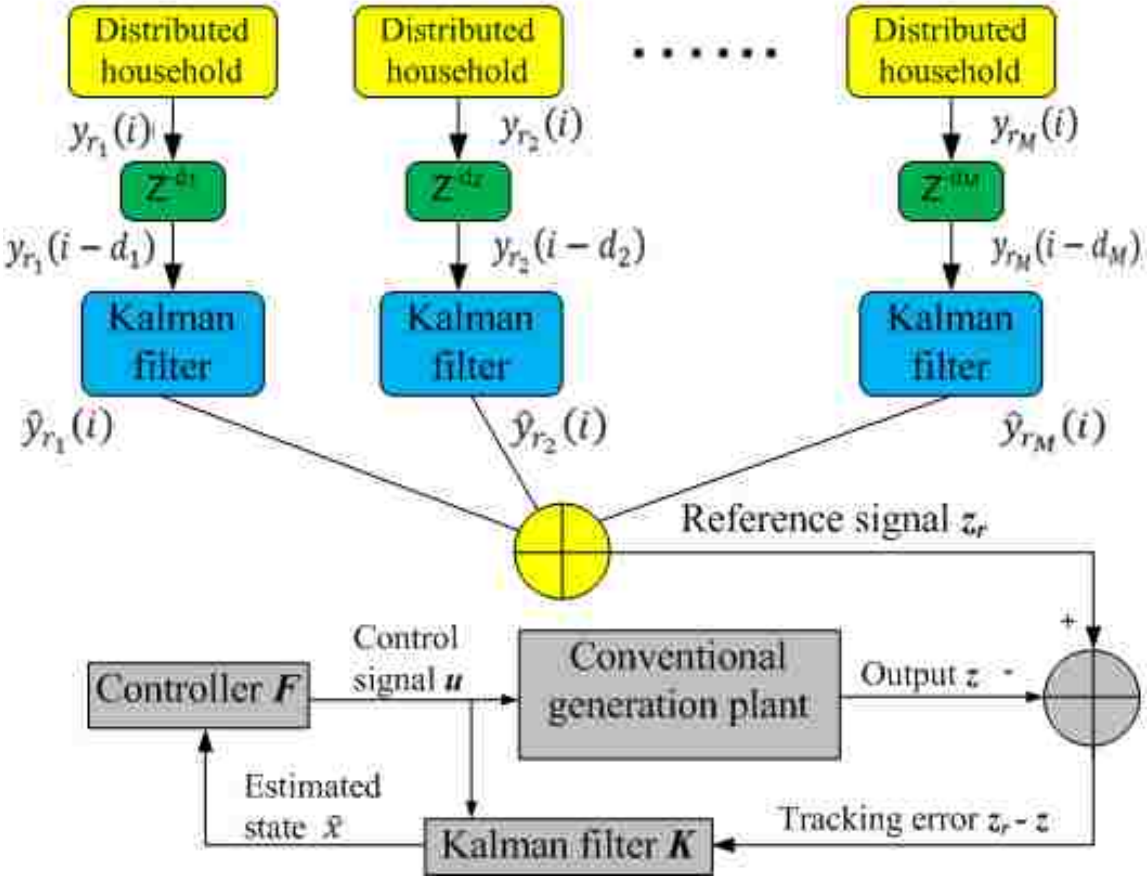

Figure 4.8: The tracking control framework with multiple asynchronous signals from different customers. Distributed Kalman filters are implemented to optimally reconstruct the non-delayed signal for each customer. The synchronous non-delayed signals are added up to construct the reference which the conventional generation needs to track.

turns out to be the sum of all reconstructed individual household demand signals. The distributed Kalman filter implementation is shown in Fig. 4.8.

The separate design in reference prediction and tracking control, which has the "prediction before control" structure is not only natural and intuitive, but also is overall optimal with respect to the original objective function (A.7). The proof is omitted due to the space limit.

4.6 Simulation Results of Reference Prediction based Tracking

We simulate the prediction and tracking schemes presented in previous sections for the assumed electricity market. The customer load demands are generated with the TVAR model. The TVAR model parameters are obtained based on the load data selected from the huge data pool of the Electric Reliability Council of Texas [53]. The dynamics of renewable generators and the conventional generators can be the same or different depending on the types of generators adopted.

With 10 customers in the electricity market, Fig. 4.9 shows the tracking performance for a scenario in which the conventional generator and the distributed renewable generators considered have different dynamics. We generate the different parameter matrices by perturbing the eigenvalues of one parameter matrix. Also, we assume that the load demand signals sent from distributed customers experience different but fixed delays. Figure 4.10 shows the prediction performance of the equivalent distributed Kalman filter implementation. It can be observed that different customers have different prediction errors due to their different system dynamics. Figure 4.11 presents the tracking error per customer as a function of the number of customers. The tracking error is approximately constant as the number of customers increases from one to ten. This matches our intuition since customers are assumed to be independent with each other.

4.7 Conclusion

An optimal stochastic tracking scheme was proposed in an interactive smart grid infrastructure. Optimal stochastic control schemes for the active power control (pri-

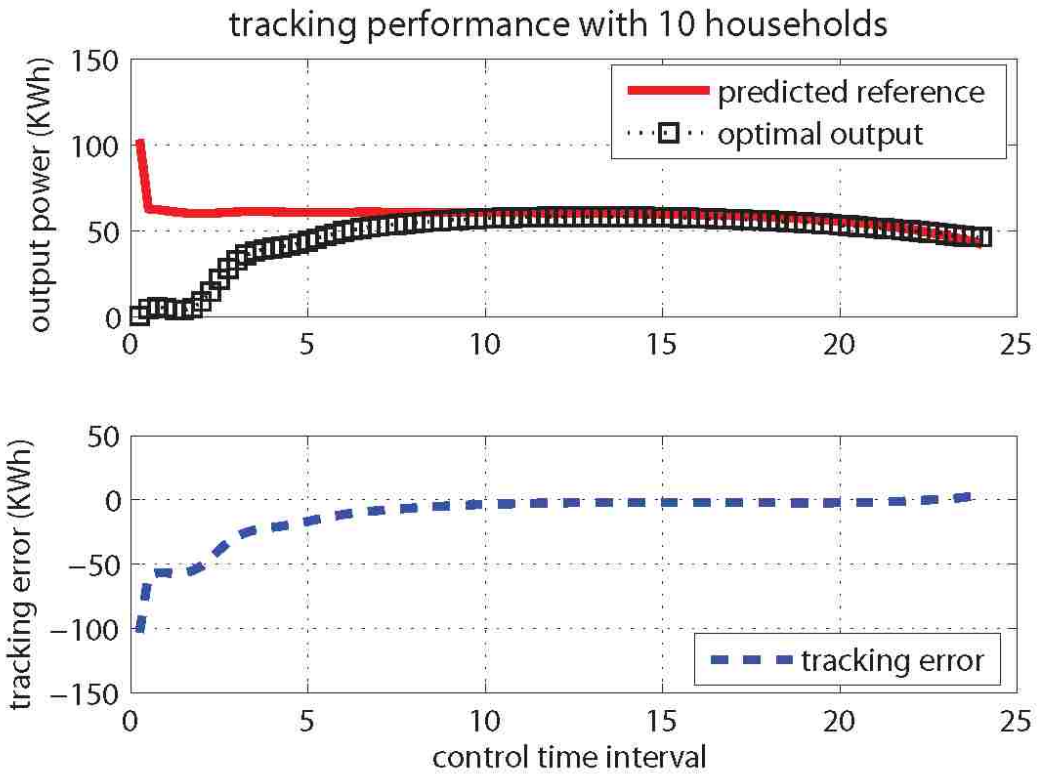

Figure 4.9: Tracking performance of the conventional power generation control. The conventional power plant and the distributed ten renewable DERs have different dynamics. Load demand signals sent from different customers experience different delays.

mary frequency control) were designed, in the presence of uncertainties arising from customer load demands and distributed renewable generations, to stabilize frequency and maintain a balance between generation and consumption within the distributed synchronous area. We proposed two stochastic tracking schemes based on the state-space representation of a synchronous generator: (1) reference dynamics-based tracking and (2) reference statistics-based tracking. We further extended the proposed optimal controllers by considering the realistic scenario of asynchronous load demand signals from different households. To compensate for different delays seen by different household signals, a Kalman filter (KF) based prediction scheme was proposed

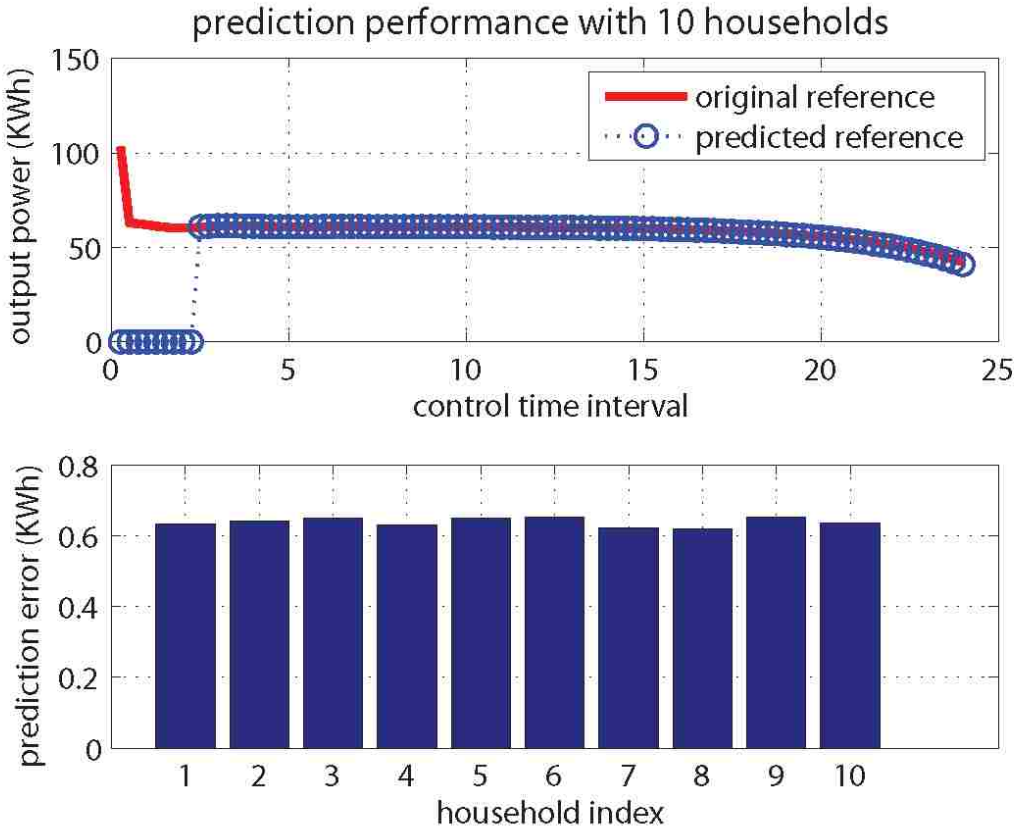

Figure 4.10: Reference prediction performance of a distributed Kalman filter implementation.

to generate the correct reference signal and we showed that the centralized reference prediction could equivalently be implemented distributively. Simulation results were presented to show the performances of the proposed prediction and tracking schemes.

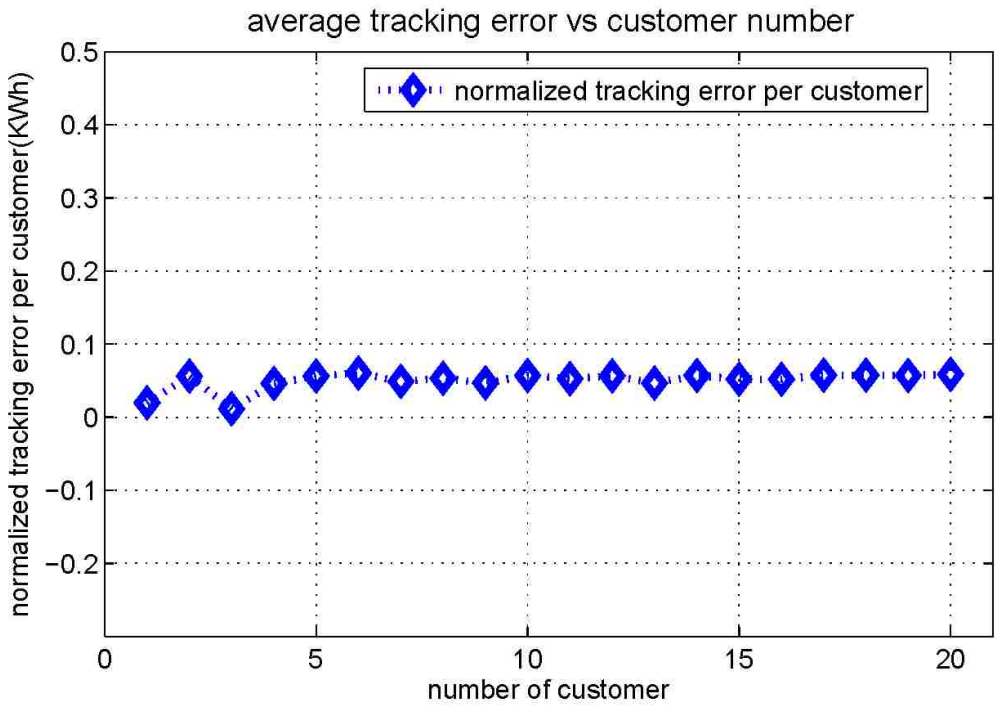

Figure 4.11: The tracking error is approximately constant as the number of customers increases from one to ten since customers are assumed to be independent with each other.

Chapter 5

Machine-learning Aided Optimal Customer Decisions

5.1 Introduction

Proliferation of distributed energy resources (DER), in particular renewable distributed generation, provides great promise in significantly improving the efficiency of electricity distribution. However, as DER's proliferate to a significant fraction of the overall electric energy on the distribution network, without proper procedures integration may lead to highly imbalanced transient behaviors which may overwhelm current infrastructure not to mention outages and brown-outs. *In a future smart grid, a customer with renewable generation capability (such as PV panels and wind turbines) may use predictive strategies to optimize its energy demand requests over time and determine when to use, sell or store its own renewable generation, flexibly interacting with the electric-grid and other customers, as opposed to being a passive energy consumer as today.* The information shared among distributed nodes (customers) endowed with generation, storage and consumption attributes can result in

a distributed decision and control framework that will lead to both overall energy and cost efficiencies. Realizing the full potential promised by smart grid concept, however, requires systematic design principles, a comprehensive protocol framework for interaction among distributed entities that make up the grid and robust and computationally efficient control and optimization algorithms.

Although a comprehensive formulation and an analysis is not yet available, still there have been some attempts to understand, model and analyze these effects [4,5]. For example, a multi-stage frequency control framework is presented in [6–8]. However, it does not address the issue of consumption planning on the customer side. The uncertainty in supply due to integrated renewable DER's and the challenges they impose on the existing distribution infrastructure and the system operator have been discussed in [9]. The distribution-level smart grid features such as interconnection of distributed generation and active distribution management, automated meter reading (AMR) systems in network management and power quality monitoring were discussed in [10]. In [11], the implementation of vehicle-to-grid (V2G) power issues, strategies and business models for doing so, for purposes of both stabilizing the grid and supporting large-scale renewable energy were discussed. Various control-theoretic and system-level problem formulations of smart grid architectures have been discussed in [12] and [13]. In [12], for example, the authors showed that significant improvements can be made to the operations of a smart grid by providing information about the likely behavior of renewable energy through both online short-term forecasting and longer-term assessments. In [13], a distributed control method was discussed for converter-interfaced renewable generation units with active filtering capability.

The topic of customer decision making consists of several important subtopics, including smart-home design [44] and [45], system integration of distributed energy resources (DER) [46], renewable generation modeling [18,19,47], load demand model-

ing [20] and [21], and plug-in hybrid electrical vehicle (PHEV) vehicle-to-grid (V2G) management and regulation [48] and [49]. There is a considerable amount of previous studies reported on these subtopics in literature. For example, a smart-home energy management system based on a ZigBee sensor network was proposed in [44]. In [45], the author motivated the use of power line LANs as a basic infrastructure for building integrated smart homes, proposing protocols capable of supporting power line communication networks at speeds comparable to wired LANs. These smart-home models are mostly from the perspective of information gathering and transmission (e.g., a ZigBee sensor network and a power line LAN). However, it is unclear how these smart-home models can be evolved to allow real-time decision making that makes use of all collected information. In [46], the concept of virtual power plant (VPP) was developed to enhance the control of DER by the system operators and other market actors by providing an appropriate interface between these system components. However, an equally important issue on the customer side (rather than the system operator side), which is the distributed self-management of DER's with local objectives, remains unaddressed. Various stochastic models for different renewable generations have been proposed in previous literature. For example, wind speed distributions are often characterized by Weibull, Rayleigh or normal distributions [17, 18]. Beta distribution has been validated by different research as a simple and sufficiently flexible two-parameter distribution to fit the empirical solar irradiation behavior data in many situations [19]. These stochastic models are important, but these papers failed to present further discussions on how these models can be incorporated in customer decisions. Similar issue arises with papers focusing on customer load modeling and prediction, for example in [20] and [21]. In [48], the impact of charging PHEVs on a distribution transformer under different charging scenarios were examined. In [49], the author established a series of well-defined electric vehicle loads that were subsequently used to analyze their electric energy usage and storage in the context of more electrified road transportation. The PHEV

management strategies mentioned above are part of the customer decision making addressed in this paper. However, it is important to consider more general energy decisions, rather than only focusing on PHEV charging strategy, taking into account of other factors, such as the impact of intermittent renewable generations. The work presented by the literature mentioned above provide an important foundation for upgrading the conventional grid-customer models to smart customers in a modern smart grid. However, little of these existing studies has considered a comprehensive cycle of interactions between the Utility and the distributed entities (customers) taking into account aspects of customer-side decision making, Utility-side demand response scheduling, renewable DER integration and power-load balances for grid-stability and the effects of information and communication technology (ICT) infrastructure on all these.

This chapter discusses the following:

1. A comprehensive architecture that addresses not only the generation control and the consumption planning *separately*, but also the interaction and integration of the two within a unified framework.

2. A hierarchical architecture, as shown in Fig. 5.1, that not only assures the scalability of the grid model, but also allows for sufficient resource pooling among customer units. This enables us to focus on an abstract power grid model consisting of one controller and multiple customer units without loss of generality.

3. An extension to the concept of "smart-home". In most current smart grid literature, the term "smart-home" is used to refer to households with "smart devices" such as Advanced Metering Infrastructure (AMI) [15, 87], which enables remote meter reading and electricity bill estimation based on real-time pricing information. In this paper, the concept of "smart-home" is extended in

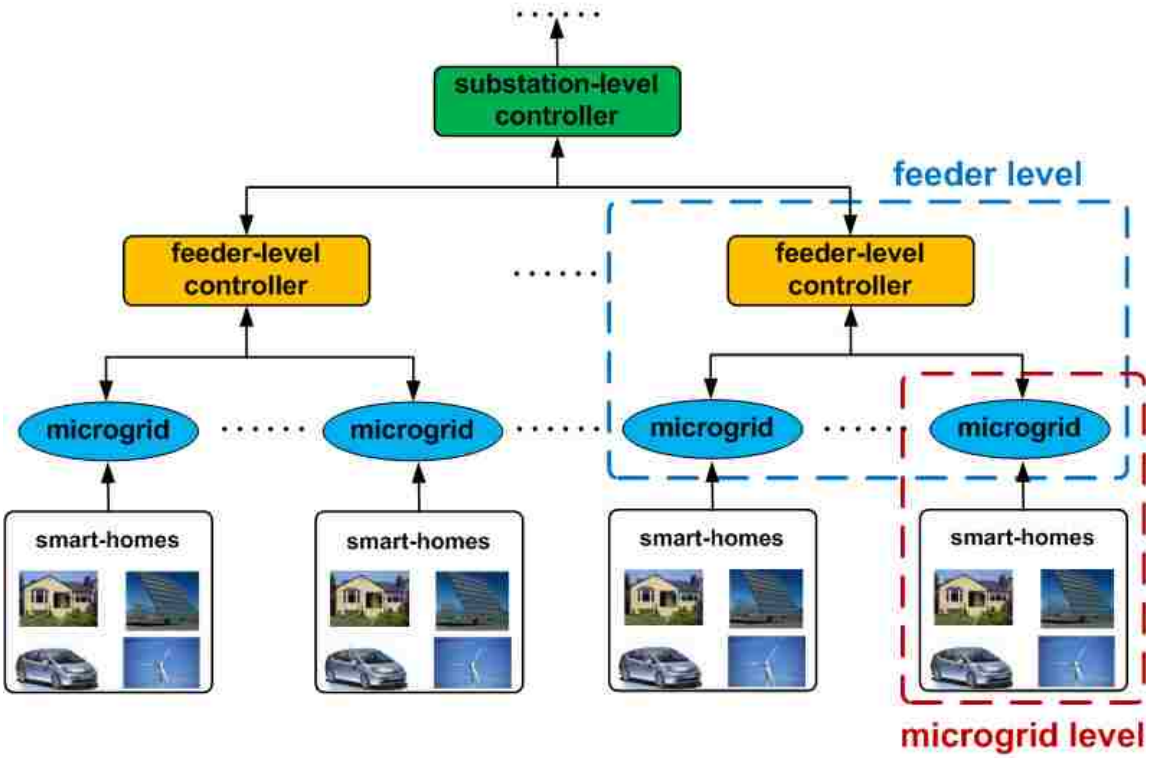

Figure 5.1: Hierarchical smart grid architecture that is scalable while allowing for sufficient resource pooling.

two aspects: First, smart-home is capable of not only intelligently managing its own energy consumption, but also actively interacting with the grid in real-time. Second, the concept of "smart-home" can scale up to a broader customer unit consisting of a cluster of households. For example, a microgrid can also be a broad smart customer unit in the feeder-level.

4. A hidden mode Markov decision process (HM-MDP) based model for the smart grid real-time planning. The HM-MDP model allows for the two-step decision framework containing both centralized sequential decision making at the controller and the auctioning game design among distributed customers.

5. A novel auctioning game for distributed customers to compete for limited en-

ergy trading opportunities. The proposed auctioning game with a reserved price has several advantages: (1) being robust to adding/removing customers, (2) being robust against collusion by customers with untruthful bidding strategies, and (3) converging to the unique Bayesian-Nash equilibrium.

It is worth pointing out that application of different auction schemes for smart grid problems have been reported in [50–52]. For example, auction mechanisms that can be used by the aggregators for procuring stochastic renewable generations are proposed in [50]. In [51] and [52], double auction is adopted for distributed energy resources (DERs) management and Plug-in hybrid electric vehicles (PHEVs), respectively. However, most of these are focused on the solution derivation of auctions and fail to address the connection between the centralized and distributed decision schemes, which is important for the hierarchical architecture of the modern smart grid.

The rest of this chapter is organized as follows: In section 5.2, we present our hierarchical interactive smart grid architecture with the proposed two-step decision-making framework. The fine-scale planning at the customer level is formulated as a sequential decision making problem in section 5.3. In section 5.4, a hidden mode Markov decision process (HM-MDP) model is investigated. In section 5.5, we analyze the non-stationary environment dynamics. A value iteration based exact solution algorithm is presented in section 5.6. In section 5.7, we further investigate a Q-learning based approximate dynamic programming (ADP) algorithm. In section 5.8, with a hidden mode Markov decision process (HM-MDP) framework well investigated for the centralized microgrid controller sequential decision making, we present a Vickrey auctioning game for the distributed customer decision making problem and the truthful bidding strategy is discussed in detail. In section 5.9, a detailed analysis on the solution set of the Bayesian Nash equilibria is presented. By introducing a reserve price, the Vickrey auction is shown to be more robust against collusive

customers and converges to the unique truthful bidding Bayesian Nash equilibrium. The conclusions from this study are given in section 5.10.

5.2 A Hierarchical Interactive Architecture

5.2.1 A Hierarchical Smart Grid Architecture

A hierarchical architecture for the smart grid that is scalable while allowing for sufficient resource pooling is shown in Fig. 5.1 [14]. The scalability of the grid requires being able to easily integrate additional customers into the grid without affecting the established operational conditions of the grid. Ideally this might be achievable if each individual household is managed separately, but, of course, this would preclude any resource pooling, which is one of the most important strategies to energy efficiency in the grid. A tradeoff to this can be achieved by using the notion of microgrids with DER's. Each microgrid is a collection of households with certain self-containing capabilities, which are geographically adjacent and coordinated by a microgrid controller, as shown in the red box in Fig. 5.1. However, we can also think of each approximately self-contained microgrid as a broader customer unit coordinated by a feeder-level controller as shown in the blue box in Fig. 5.1. Similarly, we can scale up to the substation level and above and investigate an entire hierarchical smart grid architecture, as shown in Fig. 5.1.

As we scale up to construct the entire grid, at each level, all branches with the same structure of one controller and multiple customer units are all approximately self-contained and are coordinated by the controller at a higher level. For example, at the microgrid level in Fig. 5.1, all microgrid branches identical to the red box are approximately self-contained. When the power-load mismatch is too big to be mitigated within a single microgrid, electric power flow will be routed among dif-

ferent microgrids under the coordination of a feeder-level controller. Similarly, at the feeder-level in Fig. 5.1, all branches identical to the blue box are approximately self-contained. Power flow among feeder-level branches are to be coordinated by the substation-level controller. Hence, with this hierarchical architecture interpretation, any decision-making framework designed for a controller and the individual units below it is applicable to each of the levels in this hierarchical smart grid. Thus, in the following, we may focus on an abstract model made of a single (micro-grid) controller and a collection of multiple (smart-home) customers managed by it.

It is also important to notice that this hierarchical architecture can be robust against cascading failures in a power grid due to its design based on self-containment at various granular levels. When the deviation is too large to be mitigated, the Utility can temporarily isolate the individual branch, in which the initial failure started, from the grid to prevent cascading failures. Therefore, this hierarchical structure significantly enhances the power grid reliability by routing power flow within and across different customer units to mitigate uncertainties.

5.2.2 A Utility-Customer Interaction Model between the Generation and Consumption Sides

Utility-customer (generation-consumption) interaction is an important aspect of smart grid design. The interaction between the generation and consumption sides can lead to more efficient power-load scheduling compared to conventional power grid planning, which is purely matching generation to demand. However, the Utility-customer interaction varies depending on different time scales of the interaction periods, as well as different customer units at different levels of the hierarchical architecture. For example, in a microgrid, the smart-homes are the customer units at the microgrid level while the microgrid is a customer unit at the feeder-level (one level above).

To address different interactions between the generation and consumption sides, in this section, we propose a two-stage model for the Utility-customer interaction, consisting of the initial scheduling (long-term planning) and the real-time scheduling (short-term planning).

Initial Scheduling: Prediction based Long-term Planning

In the initial interaction (long-term planning) stage, demand response (DR) schemes are implemented and it is desired that the customer loads always stay relatively flat. Note that, a flat load profile with low peak-to-average ratio means a need for relatively low generation capacity reserve, leading to more efficient operations of conventional generation facilities and a less number of idle generators for most of the time. Usually, peak load shaving and load profile flattening can be achieved by incorporating demand response (DR) schemes that are based on the predicted renewable generation. Various demand response (DR) schemes have been reported in literature [30,60,88–92] based on different pricing schemes such as time-of-use (TOU), peak-time pricing and real-time pricing [93]. In [30, 31], we presented optimization-based and game theoretic DR schemes for the Utility to achieve this goal. In these DR schemes, customers pay less (or receive incentive payments) if they strictly fulfill their energy commitments. Similarly, they will have to pay extra as a penalty if they fail to honor the agreement reached during the long-term planning. However, DR schemes only provide a nominal operating point for the nodes in the grid (i.e. the flat load profiles for customers) without allowing for the real-time fluctuations and intermittence in the grid due to the inevitable mismatches between the actual and predicted renewable generations. Interaction at this level usually happens at the beginning of each scheduling period [30, 60, 91, 92] and is called the initial interaction or long-term planning in our interaction framework.

Real-time Scheduling: Short-term Planning

The DR schemes in the initial scheduling provide a nominal operating point for the nodes in the grid (the flat customer load profiles). However, since all DR schemes are based on the prediction of the renewable generations within the scheduling period (for example, a 12-hour or 24-hour period), they may not properly handle the real-time fluctuations and intermittence in power grid due to the inevitable mismatches between the actual and predicted renewable generations. This can be overcome, and the overall efficiency and stability can be improved, by allowing for (near real-time) interactions at a finer time scale (short-term) between the Utility and customers (generation and consumption sides).

From the perspective of the Utility (conventional generation side), both frequency control and voltage control schemes are needed for keeping active and reactive power-load balances [6,7,16,73]. From the perspective of customers (consumption side), who are most likely self-oriented, the objective is to make optimal decisions to maximize the accumulated profits (or minimize the payments) by taking advantage of their local DER's. Given the relatively flat load profiles computed by DR schemes, a customer can decide to sell part of its excess renewable energy to the grid and storing the rest for future use, according to the real-time pricing information.

It is worth pointing out that though the real-time decision schemes are important supplements to the DR schemes, they are different not only in scales of scheduling period but also in their functionalities. The DR scheme design (long-term planning) provides a nominal operating point (flat load profiles) for the nodes. In real-time scheduling, on the other hand, if local generations are less than the nominal load demands computed in the long-term planning, customers do not have much flexibility other than to buy electricity they need from the Utility. However, if local renewable generations are more than the nominal load demands, customers can flexibly decide

how much of their own excess energy to be sold. Therefore, no matter in what scenario (buying or selling), the real-time scheduling is always based on the flat load profiles computed by the DR schemes.

5.2.3 A Two-step Decision Framework for Real-time Scheduling

Various DR schemes for the long-term planning problem have already been reported in literature [30, 60, 88–92]. In this chapter, however, our focus is on the real-time scheduling problem in the above assumed abstract grid model (consisting of one controller and multiple customer units). Within this abstract model, there are two main decision problems (for real-time scheduling) to be addressed: (1) centralized controller decisions and (2) distributed customer decisions.

Take a microgrid as an example. On one hand, as a customer unit at the feeder-level, the microgrid controller needs to make sequential decisions to maximize the accumulated reward of the entire microgrid. At each time step the microgrid controller decides how much electric energy need to buy or sold by the microgrid, taking into account of all local DER's within the microgrid (first problem). On the other hand, smart-homes (customers) with excess energy also need to make distributed decisions when the microgrid controller needs to sell part of the excess energy. The distributed decisions indicate how much excess energy each smart-home contributes to the total amount of electric energy to be sold by the whole microgrid (second problem). To address both the centralized and distributed decision making problems, we propose a two-step decision framework for real-time scheduling, as shown in Fig. 5.2. The centralized microgrid controller decision making problem is shown in the upper level in Fig. 5.2 and the distributed smart-home decision making problem is shown in the lower level in Fig. 5.2. In light of the discussion on how the abstract model

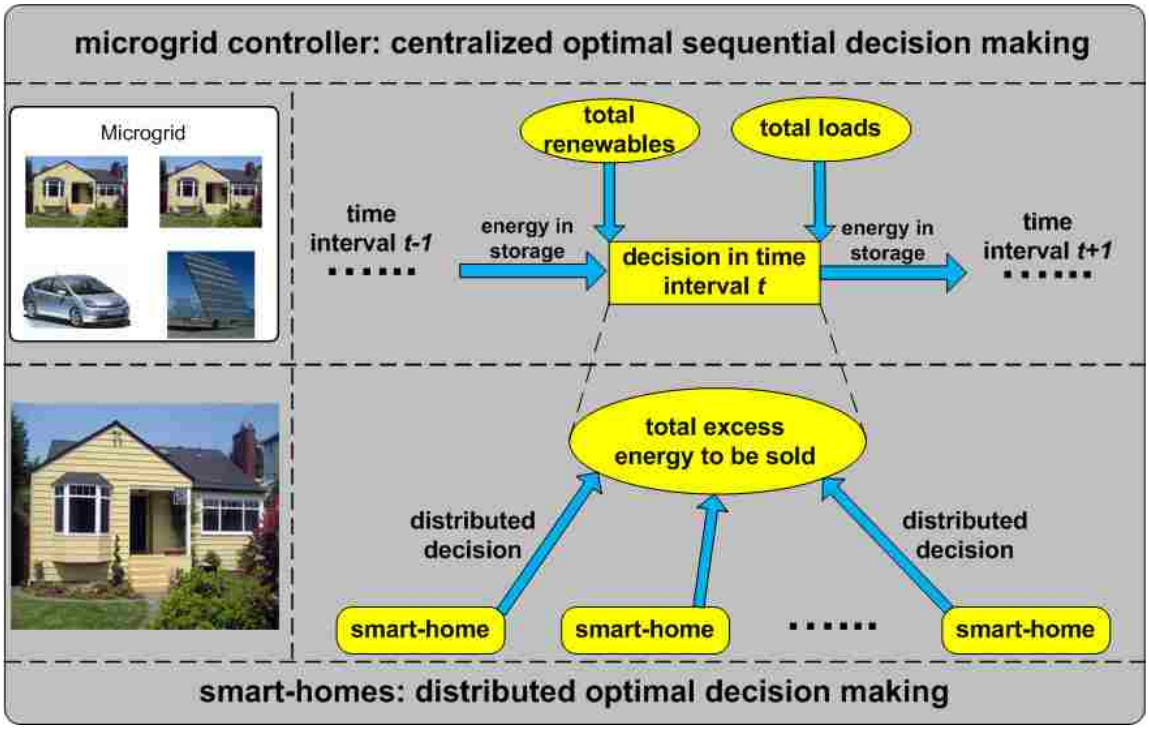

Figure 5.2: A two-step decision framework for a microgrid addressing (1) centralized microgrid controller decisions and (2) distributed smart-home decisions.

can represent scaled up units in the hierarchical model, the optimal decision making strategies developed for this abstract model can also be applied to different levels in the hierarchical smart grid with relevant modifications.

5.3 Centralized Sequential Decision Making at Microgrid Controller: Problem Formulation

The sequential decision making problem faced by a microgrid controller is illustrated in Fig. 5.3. The objective is to maximize the accumulated profit of the entire microgrid (consisting of multiple smart-home customers) over a predefined scheduling

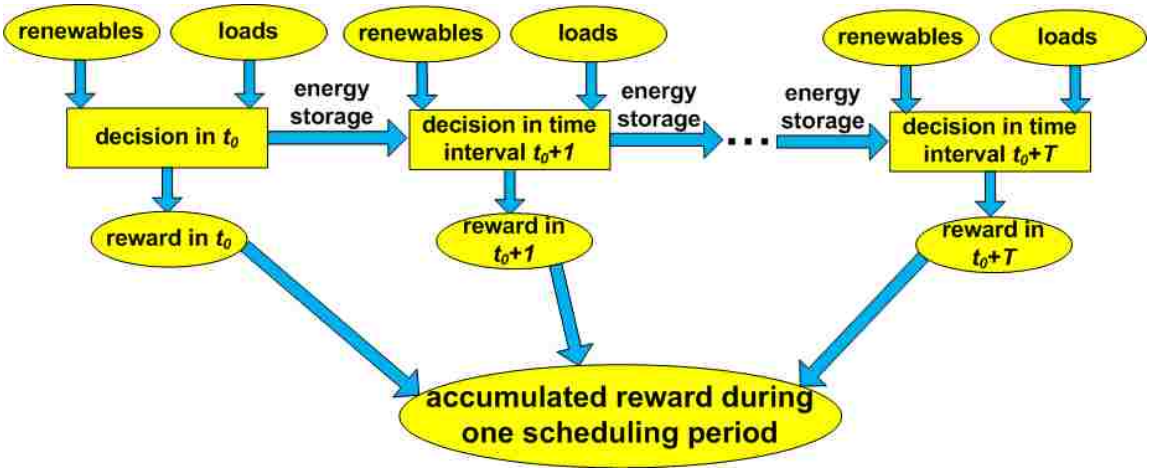

Figure 5.3: Microgrid controller makes optimal sequential decision to maximize the total accumulated reward over the scheduling period.

period of length T. In time interval t ($1 \leq t \leq T$), the microgrid controller takes into account of all available local information, such as energy storage, renewable generation and load demands within the microgrid. Depending on whether the stored energy and local renewable generation are enough to support the total load demands, the microgrid controller either decides to sell part of its total excess energy to the grid or buys electricity it needs from the Utility, taking into account the possibly time variant pricing information.

Without loss of generality, for an arbitrary microgrid controller, we denote by $L(t) \in [L^{min}, L^{max}]$, $E_r(t) \in [E_r^{min}, E_r^{max}]$, and $E_s(t) \in [0, E_s^{max}]$ the total load demand, renewable energy generation and stored energy, respectively, of this microgrid in time interval t. The renewable generation $E_r(t)$ and the load demand $L(t)$ are both stochastic in nature [17,18,94]. We further define the excess energy in time interval t as $E_x(t) = E_r(t) + E_s(t) - L(t)$, where $E_x(t) \in [E_r^{min} - L^{max}, E_r^{max} + E_s^{max} - L^{min}]$. Note that negative excess energy implies buying energy from the Utility.

The Markov properties of load demands and renewable generations [95–97] enable

us to characterize the transitions of the quantities defined above by a Markov decision process (MDP) (possibly non-stationary). Denote by $s(t)$ the state of the MDP, which is defined as a two dimensional vector $s(t) = [E_s(t), E_r(t) - L(t)]$. The state space is denoted by

$$s_1(t) = E_s(t), \quad s_2(t) = E_r(t) - L(t), \tag{5.1}$$

where $s_1(t) \in \mathbb{S}_1 = [0, E_s^{max}]$ and $s_2(t) \in \mathbb{S}_2 = [E_r^{min} - L^{max}, E_r^{max} - L^{min}]$. Thus we have

$$E_x(t) = s_1(t) + s_2(t). \tag{5.2}$$

Though $E_r(t)$, $E_s(t)$ and $L(t)$ are continuous-valued quantities, we can quantize the state space into discrete levels. On one hand, a certain level of granularity, say, a "basic energy unit", is essential in practice for energy operations to be effective enough for planning at different levels. On the other hand, the error can be made sufficiently small by making the quantization level as small as desired. It is worth pointing out that power scheduling at different levels of the hierarchical architecture have different scales. Hence, the basic energy unit could be different for customer units at different levels. For example, the basic energy unit for a single household could be relatively small compared to that of a microgrid as a customer unit. Similarly, in large scale grid scheduling (i.e. microgrids, feeder-level grids and substation-level grids), the basic energy unit could be relatively large, helping to reduce the size of the discrete-state space $\mathbb{S}$.

As mentioned earlier, positive excess energy of a microgrid can be fully or partly stored for its own use in future or be sold to the Utility, depending on the specific decisions of the microgrid controller. Depending on the sign (+ or −) of $E_x(t)$, a

microgrid switches between two operating modes: the *buying-mode* and the *selling-mode*. A microgrid is said to be in buying-mode when $E_x(t)$ is negative and in selling-mode when $E_x(t)$ is positive. For a microgrid in buying-mode, the only available action is to buy the energy it needs from the Utility. However, for a microgrid in the selling-mode, we can define a decision variable $a(s(t), t)$ for the selling-mode microgrid controller as the number of basic energy units to be sold to the Utility, where the action $a(s(t), t)$ is not only a function of time but also a function of the current state $s(t)$. Note that, though the maximum value for $a(s(t), t)$ is $E_x(t)$ and therefore time varying, we can still define an overall time invariant action space as

$$\mathbb{A} = \left[0, 1, 2, \ldots, E_r^{max} + E_s^{max} - L^{min}\right], \tag{5.3}$$

where $E_r^{max} + E_s^{max} - L^{min}$ is assumed to be an integer multiple of the basic energy unit. Then in time interval $t+1$, the available energy from storage is given by

$$\begin{aligned} s_1(t+1) &= E_s(t+1) \\ &= \max\left[0, E_x(t) - a(s(t), t)\right] \\ &= \max\left[0, s_1(t) + s_2(t) - a(s(t), t)\right]. \end{aligned} \tag{5.4}$$

This MDP model is shown in Fig. 5.4 and the corresponding time variant state transition function is given by (5.5).

$$\begin{aligned} s(t+1) &= \left[s_1(t+1), s_2(t+1)\right], \\ &= \left[\max\left(0, \min\left(s(t) \cdot [1,1]^T - a(s(t), t), E_s^{max}\right)\right), E_r(t+1) - L(t+1)\right] \end{aligned} \tag{5.5}$$

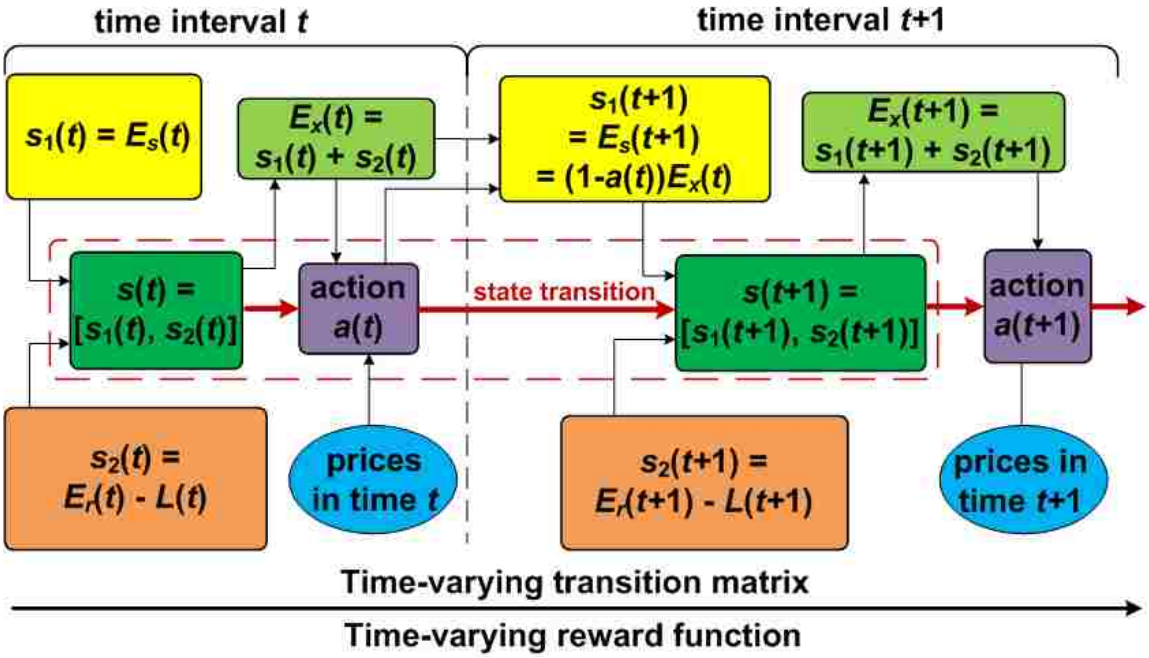

Figure 5.4: The non-stationary Markov decision process (MDP) model for the centralized controller decision making problem.

Given the distribution information of the random quantity $s_2(t+1) = E_r(t+1) - L(t+1)$ [94], the state transition with probability $P^a_{ss'}$ of this MDP is defined as a mapping from the set of current states and actions to the state space, $\mathbb{S} \times \mathbb{A} \rightarrow \mathbb{S}$. For each time interval, in addition to the local information we defined above, the microgrid controller also receives pricing information from the Utility for the current and future time intervals within the scheduling period. In general, there are two types of prices: the *buying price* $p_b(t)$ and the *selling price* $p_s(t)$. The buying price is defined as the price at which the Utility buys energy from the microgrid. The selling price is defined as the price at which the Utility sells energy to the microgrid. As the fixed pricing scheme is being replaced by dynamic pricing schemes in practice [31], both buying and selling prices can be time varying. In this chapter, we make following important assumptions:

1. Microgrids are allowed to purchase electric energy only for consumption and not for investment. Therefore, microgrids buy electric energy from the Utility only when their local renewable generations are less than their total load demands.

2. Microgrids are allowed to store only their own excess local renewable generations. Therefore, microgrids will store energy only when they have excess local renewable generations after supporting their own total load demands.

3. Microgrids are allowed to trade electric energy only with the Utility. Direct trading among microgrids without coordination of a higher level (feeder-level) controller is not allowed.

It is worth pointing out that, the third assumption that "energy trading among microgrids is not allowed" does not mean "energy flow among microgrids is not allowed". Indeed, appropriate energy routing among different branches of the entire grid is important to ensure resource pooling and the stability of the entire grid. Selling excess electric energy back to the Utility does not mean transmitting all excess power back to a certain "center point" and then redistributing over the entire grid. The role of the Utility is to make sure that energy routing among customer units (e.g., microgrids) is under the coordination of a higher level controller.

$$
\begin{aligned}
R_t &= \begin{cases} p_b(t)a(s(t),t) & \text{if } s_2(t) \geq 0 \quad , \\ p_b(t)a(s(t),t) + p_s(t)|s_2(t)| & \text{if } s_2(t) < 0, s_1(t) + s_2(t) \geq 0 \quad , \\ p_s(t)s_1(t) & \text{if } s_1(t) + s_2(t) < 0 \end{cases} \\
&= p_b(t)a(s(t),t)\mathbb{I}_{\{s_2(t)\geq 0\}} + p_b(t)a(s(t),t) \\
&\quad + p_s(t)|s_2(t)|\mathbb{I}_{\{s_2(t)<0, s_1(t)+s_2(t)\geq 0\}} + p_s(t)s_1(t)\mathbb{I}_{\{s_1(t)+s_2(t)<0\}}. \qquad (5.6)
\end{aligned}
$$

Given the pricing information defined above, the reward function $R_t(s_t, a(s(t),t))$

of the microgrid in time interval t can be defined as in (5.6), where $\mathbb{I}_{(\cdot)}$ is the indicator function defined as in

$$\mathbb{I}_{\mathbb{E}} = \begin{cases} 1 & \text{if event } \mathbb{E} \text{ is true} \\ 0 & \text{if event } \mathbb{E} \text{ is false} \quad . \end{cases} \tag{5.7}$$

The first two terms in (5.6) correspond to the selling mode and the third term corresponds to the buying mode. Generally, the reward is defined as the sum of two parts:

1. How much the microgrid has gained by selling part of the excess energy;
2. How much the microgrid has saved by consuming the stored energy compared to buying the same amount of energy from the Utility.

Note that the reward is 0 when the microgrid needs to buy energy from the Utility. This is because that reward is a function of the action and when the microgrid needs more electric energy, there is no other action but to buy the required electricity from the Utility. The microgrid can really make active decisions only when there are excess energy to be sold. When there is no choice (or unique choice), it makes sense the reward to be a constant value (which is zero). In particular, negative rewards will be inconsistent with this problem formulation.

A *decision rule* is a mapping from the set of states to the set of actions, $d_t : \mathbb{S} \rightarrow \mathbb{A}$ [98]. A *policy* π is a sequence of decision rules $\pi = (d_1, d_2, \ldots, d_T)$ for the entire planning period, where d_t is the decision rule for the time step t. For each policy we define a value function (expected return) which is the expected discounted sum (with discount factor γ, $0 < \gamma \leq 1$) of rewards of the microgrid over the scheduling period given the current state s:

$$\tilde{V}_t^{\pi}(s) = \mathbb{E}_{\pi}\left\{\sum_{\tau=0}^{T-1}\gamma^{\tau}R_{t+\tau}(s_{t+\tau}, a_{t+\tau})|s_t = s\right\} \tag{5.8}$$

The objective of each microgrid is to find an optimal policy π^* that maximizes the value function:

$$\pi^* = \underset{\pi}{\operatorname{argmax}}\ \tilde{V}^{\pi} \tag{5.9}$$

Let us denote by $\tilde{V}^* = \tilde{V}^{\pi^*}$ the optimal value function. For a finite horizon scheduling, i.e. with finite T , the optimal policy is usually non-stationary because the "number of steps to go before termination" is different at each time interval.

5.4 A Hidden Mode Markov Decision Process (HM-MDP) Model for Centralized Decision Making

In the above MDP model (5.5) for microgrid sequential decision making problem, we naturally include all local information, such as $L(t)$, $E_r(t)$ and $E_s(t)$, into the state of a microgrid. As a result, this proposed MDP is difficult to solve due to its non-stationarity: Both the state transition probability matrix and the reward function are time varying. The non-stationary dynamics come from the fact that both load demand $L(t)$ and renewable generation $E_r(t)$ are, in general, non-stationary stochastic processes. For example, the load demand during a day can change dramatically from peak hours (afternoon and evening) to non-peak hours (midnight and early in the morning) [97] and the renewable generation can be heavily dependent on weather conditions [73, 94, 99].

In addition to various renewable generation models (Weibull, Rayleigh and Beta distributions) proposed in literature mentioned earlier, stochastic approaches have also been reported in the literature for household load modeling and two approaches are widely adopted. The first approach is component-based load modeling approach, which reconstructs the expected daily electrical loads of a household based on appliance sets, occupancy patterns, and statistical data. For example, in [20], the authors constructed such electric load profiles from individual appliance profiles. By considering *availability* and *proclivity* functions, they predict whether someone is available (at home and awake) and their tendency to use an appliance at any given time. These functions were applied to predict individual appliance events, which were then aggregated into a load profile. The second approach is termed the measurement-based load modeling. In [21], the authors used this approach to create electrical profiles to examine demand side management strategies for Finland. In [22], a methodology of measurement-based load modeling for transient stability analysis was proposed and Genetic Algorithms (GA) was used to estimate load model parameters. In [97] we also proposed a time series technique based model to describe the randomness in load demand. Moreover, the dynamic pricing information provided by the Utility can also be time varying.

In the following we propose to transform the above non-stationary MDP into a stationary hidden mode MDP so that efficient solution algorithms can be developed.

5.4.1 Hidden Modes and State Transitions in a Non-stationary Environment

Generally, it is very difficult to solve MDP problems with arbitrarily changing dynamics [99]. Most standard MDP studies assume stationary dynamics (i.e., MDP transition matrix P and reward function R are time invariant). Solution methods for

non-stationary MDP problems are usually obtained only under certain assumptions. With more information about how the dynamics change, better modeling and solutions can be expected. One important class of algorithms for MDP with unknown and time variant environment are model-based [100]: For example, they compute policies by first estimating transition and reward functions and then solving the estimated MDP. Model based approaches do not need too much specific information about how the dynamics change. However, these algorithms are particularly vulnerable to fast-changing dynamics, since they typically employ maximum likelihood estimates of parameters of the model, based on the past experience. Thus, it is very likely to average into a large body of outdated prior data. If customers do not know how the dynamics change but know the rate of changes, an on-line algorithm can be employed to keep track of the changes by giving higher weights on the near history [101]. Other approaches for solving MDP with non-stationary dynamics have also been reported in literature: In [102], the model uncertainty is represented by a Dirichlet distribution over possible models. Parameters of the distribution are updated directly as new experience is acquired. To cope with non-stationarity, the parameters are decayed as time goes. In [103], the non-stationarity is handled by assuming arbitrary but bounded variations in the transition probabilities. In [104], the change of dynamics is assumed to be always confined to a small number of choices, named environment modes, that evolves itself over time according to a Markov chain. MDP's specified by different modes have different transition matrices and reward functions but share the same state and action spaces.

In our problem, it is intuitive that we include all local information ($L(t)$, $E_r(t)$ and $E_x(t)$) in the state, but it is incomplete because there are other factors that affect the decisions of the microgrid controller, that are not considered as part of the state. These factors include weather conditions and various events that influence either renewable generations or customers' energy consumption behaviors. These external factors are not local information for microgrid controller and might not be

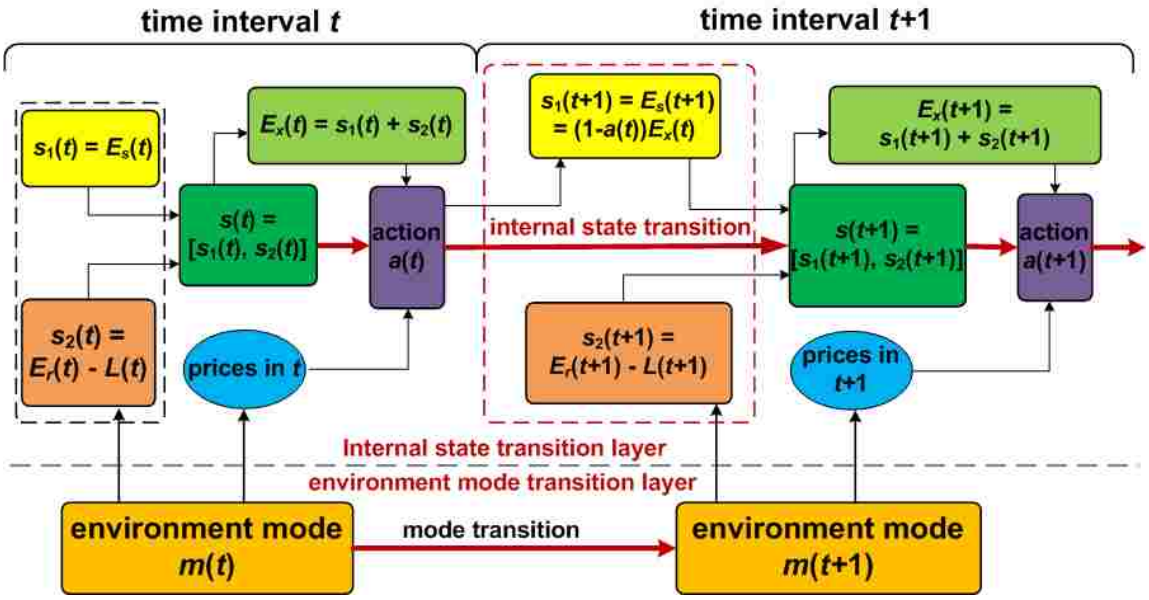

Figure 5.5: The hidden mode Markov decision process (HM-MDP) model for the centralized controller sequential decision making problem.

as fully observable as the local information is. However, they do provide us with the possibility to convert the original non-stationary MDP model to a stationary one. This may allow for the use of more efficient solution algorithms rather than pure model-based approaches which are vulnerable to fast-changing dynamics [100]. To incorporate these external factors, we define a generalized state consisting of two parts [104]: the *internal state* which is the original state as defined earlier and the *environment* which contains all external factors. It is worth pointing out that one difficulty in this state generalization is in quantizing the environmental factors. There are many factors coupled with each other in the environment making it even difficult to just determine the number of dimensions of the representation vector space [18].

However, observations on environmental transitions reveal a possible environment characterization with the concept of an *environment mode*. Since Markov models have been widely used in weather forecasting and household load prediction [95–97], we can reasonably assume that the environment can be characterized by a set of

environment modes, denoted by $m(t)$. The transition behavior of the environment among these modes can be characterized by a Markov model. Moreover, since the Utility determines the real-time electricity prices according to the total load demands and renewable generations, both buying and selling prices are actually functions of the environment mode. Unlike the state $s(t)$, which is based only on local information and fully observable by the microgrid controller, the environment mode $m(t)$ is hidden from microgrids (as it is not local information of microgrids) and can be only estimated based on observations of other information, such as renewable generations and energy consumptions. As a result, the environment mode transition can be characterized by a Hidden Markov Model (HMM) [105].

5.4.2 A Hidden Mode Markov Decision Process (HM-MDP) Model for Microgrid Controller Decision-making

With the fully observable internal state and the hidden environment mode as defined above, we may adopt a hidden mode Markov decision process (HM-MDP) model to solve the microgrid controller centralized sequential decision making problem [104–106]. In addition to the tuple $\{\mathbb{S}, \mathbb{A}, P, R\}$, now we also define the environment mode space denoted by $\mathbb{M}$. As shown in Fig. 5.5, the proposed HM-MDP transition consists of two layers: the microgrid layer (internal state transition layer) and the environment layer (environment mode transition layer).

Assume that at any time t, the environment mode is m and the microgrid state is s. Based on the observable state s the microgrid controller makes a decision a and receives an immediate reward $R^m(s, a)$. Then at time $t + 1$, the environment mode changes from m to $m^{'}$ with a transition probability of $x_{mm'}$ and the microgrid state changes from s to $s^{'}$ with a transition probability of $P_{ss'}^{a,m'}$. Note that the immediate reward function $R^m(s, a)$ is indexed by mode m because the electricity

prices (both the selling price and the buying price) can be functions of environment mode m, as the Utility usually takes these external factors into account to determine the prices [93]. Similarly, the state transition probability $P_{ss'}^{a,m'}$ is indexed by the successor mode m'. This is because the transition probability is determined by the random quantity $(E_r(t+1)-L(t+1))$ which can be a function of the successor mode m' at time $t+1$.

Note that the transition of environment mode is independent from the internal state, but the transition of internal state is dependent on the corresponding successor environment mode. This is seen by noting that the probability distribution for the random quantity $(s_2(t) = E_r(t) - L(t))$ depends on the environment mode. This probability distribution affects the distribution of the internal state transition. Similarly, $s_1(t) = E_s(t)$ is dependent on the previous environment mode through the state in previous step $s(t-1)$. From Fig. 5.5 we can see that the probability distribution of $(s_2(t))$ actually determines how the HM-MDP process evolves over time:

1. The current energy storage $s_1(t) = E_s(t)$ and the probability distribution of $(s_2(t))$ corresponding to the current environment mode $m(t)$ determines the probability distribution of current state $s(t)$.

2. The current excess energy $E_x(t)$, current decision $a(t)$ and the probability distribution of $(s_2(t+1))$ corresponding to the next environment mode $m(t+1)$ determines the transition probability of the internal state from $s(t)$ to $s(t+1)$.

5.5 Partially Observable Environment and Belief Mode Estimation

Since the environment mode is hidden from the microgrid, we define the *belief mode* in a similar way the *belief state* is defined for partially observable Markov decision processes (POMDP) [98]. The belief mode μ, which is a probability distribution over the environment modes, is a *sufficient statistic* for the optimal decision making in an HM-MDP [98]. The set of all possible belief modes is referred to as the *belief mode space* denoted by $\mathbb{B}$. We then define a decision policy for an HM-MDP as a mapping from $\mathbb{S} \times \mathbb{B}$ to $\mathbb{A}$, that prescribes an action for each pair of state and belief mode.

Knowing the transition matrix of the HMM of the environment mode is a prerequisite for calculating the belief mode transition rule. Therefore, there are two steps to calculating the belief mode:

1. HMM model learning of the environment mode;
2. Estimation of the new belief mode. Based on the observable state and the estimated belief mode, microgrid controller then make sequential decisions to maximize the accumulated rewards.

In general, the transition probabilities of the Markov chain developed for the environment mode modeling, though fixed, are unknown in advance. There are usually two ways to obtain this transition matrix. One approach is based on the parameter estimation of the Markov model of environment uncertainties [94]. Thus the accuracy of the transition matrix estimation depends on how well these uncertainties are modeled. The other approach, which is more robust against modeling error, is based on online learning, in which the current environment mode is inferred by observing the state transition history. Baum-Welch algorithm, which is adopted in this

chapter, represents this second approach [105], in which the mode HMM transition probability $x_{mm'}$ is estimated based on maximum likelihood (ML) criterion. In the following, we focus on the second step: Bayesian estimation of the belief mode.

5.5.1 HMM Transition Probability Learning: Baum-Welch Algorithm

In general, the transition probabilities of the Markov chain corresponding to the environment mode, though fixed, are unknown in advance. There are usually two ways to obtain this transition matrix. One approach is based on the parameter estimation of the Markov model of environment uncertainties [94]. Thus the accuracy of the transition matrix estimation depends on how well these uncertainties are modeled. Another approach, which is more robust against modeling error, is based on online learning, in which the current environment mode is inferred by observing the state transition history. Baum-Welch algorithm represents this second approach [105], in which the mode HMM transition probability $x_{mm'}$ is estimated based on maximum likelihood criterion, i.e. find the values of model parameters that maximize the likelihood of the observed data. It has been shown in [105] that the iterative parameter estimation procedure in Baum-Welch algorithm either increases the likelihood function or leave it constant. In the latter case, then parameter set is a fixed point [105]. Due to space limit, we omit the detailed steps of the Baum-Welch algorithm here. Interested readers are referred to [104, 105].

5.5.2 Bayesian Estimation of the Belief Mode

Assume that from time t to $t+1$, the environment mode changes from the current mode m to another hidden mode m' independent of the current and next internal

state transitions. When an action a is executed by the microgrid controller at time t, the state changes from the current state s to another observable state $s^{'}$. We denote by $\mu^{a}_{s^{'}}$ the next belief mode determined by the action a and the next state $s^{'}$ given the current state s and the current belief mode μ. Using the Bayes' formula, the $m^{'}$-th component of the next belief mode $\mu^{a}_{s^{'}}$ is given by

$$\mu^{a}_{s^{'}}(m^{'}) = \frac{Prob(s^{'}, m^{'}|s, \mu, a)}{\sum_{m^{''}} Prob(s^{'}, m^{''}|s, \mu, a)} = \frac{\sum_{m} x_{mm^{'}} P^{a,m^{'}}_{ss^{'}} \mu(m)}{\sum_{m^{''}} \sum_{m} x_{mm^{''}} P^{a,m^{''}}_{ss^{'}} \mu(m)} \tag{5.10}$$

where $P^{a,m^{'}}_{ss^{'}}$ is the conditional probability that the internal state transfers to $s^{'}$ given that the current state is s and the microgrid controller takes action a and environment mode changes to $m^{'}$ and can be calculated empirically based on real measured data.[1]

Note that in deriving (5.10) we have used the facts that: (1) The transition of environment mode from m to $m^{'}$ (with probability $x_{mm^{'}}$) does not depend on the current state s and action a, so that $Prob(m^{'}|s, m, a) = Prob(m^{'}|m) = x_{mm^{'}}$; (2) The transition of state from s to $s^{'}$ (with probability $P^{a,m^{'}}_{ss^{'}}$) does not depend on the current mode m, $Prob(s^{'}|s, m, a, m^{'}) = Prob(s^{'}|s, a, m^{'}) = P^{a,m^{'}}_{ss^{'}}$. This is because the transition of environment mode is determined by external factors such as weather conditions, which are independent from the microgrid state and the microgrid controller action. However, the transition of belief mode and the resultant belief mode do depend on the resultant state $s^{'}$. This is because the belief mode is an estimate of the hidden environment mode, which is made by the microgrid controller based on the observation of its own state (local information).

[1]Strictly speaking, we need to consider the case that the state transition from s to $s^{'}$ happens with a probability of 0, meaning $P^{a,m^{'}}_{ss^{'}} = 0, \forall m \in \mathbb{M}$ and the denominator in (5.10) is 0. In this case, the term $P^{a,m^{'}}_{ss^{'}}$ is independent of m and (5.10) reduces to $\mu^{a}_{s^{'}}(m^{'}) = \frac{\sum_{m} x_{mm^{'}} \mu(m)}{\sum_{m^{''},m} x_{mm^{''}} \mu(m)}$.

5.5.3 Maximum Likelihood Estimation of Transition Matrix $P_{ss'}^{a,m'}$

Recall that $P_{ss'}^{a,m'}$ is the conditional probability that the internal state transfers to s' given that the current state is s and the microgrid controller takes action a and environment mode changes to m'. This can be calculated empirically by using the standard maximum likelihood estimation (MLE) based on measured data. Denote by $\{p_{ss'}|0 \leq s, s' \leq l\}$ the entries of transition matrix $P_{ss'}^{a,m'}$ for arbitrary mode m' and action a. Denote by S_i the state in step i, where $p_{ss'}$ is the transition probability from state s to state s' and l is the number of states. For any pair of initial and final states s_1 and s_j, the likelihood is given by $L(p) = Pr(S_1 = s_1)\prod_{i=2}^{j} p_{s_{i-1}s_i}$. Define the transition counts $n_{ss'}$ as the number of times that state s is followed by state s'. Then we may rewrite the maximum likelihood estimation problem as

$$\underset{p_{ss'}}{\text{maximize}} \qquad Pr(S_1 = s_1)\prod_{s=1}^{l}\prod_{s'=1}^{l} p_{ss'}^{n_{ss'}} \tag{5.11}$$

$$\text{subject to} \qquad \sum_{s'} p_{ss'} = 1, \quad s = 1, 2, \ldots, l \tag{5.12}$$

Notice that the optimal estimation $\hat{p}_{ss'}$ that maximizes the log-likelihood $\log(L(p))$ also maximizes the likelihood function, where

$$\log(L(p)) = \log(Pr(S_1 = s_1)) + \sum_{s,s'} n_{ss'} \log(p_{ss'}) \tag{5.13}$$

Hence, this convex optimization problem can be solved by introducing a new objective function $\mathbb{L}(p) = \log(L(p)) - \sum_s^l \lambda_s(\sum_{s'} p_{ss'} - 1)$, where $\lambda_1, \lambda_2, \ldots, \lambda_l$ are Lagrange multipliers. Taking into account both zero derivative conditions $\partial\mathbb{L}(p)/\partial p_{ss'} =$

0 and the probability transition matrix constraints $\sum_{s'} p_{ss'} = 1$, for $s = 1, 2, \ldots, l$, we have $n_{ss'}/\hat{p}_{ss'} - \lambda s = 0$ and $\sum_{s'=1}^{l} n_{ss'}/\lambda_s = 1$ for all s. Thus, the MLE estimator of transition matrix is given by

$$\hat{p}_{ss'} = n_{ss'} / \sum_{s'} n_{ss'} \tag{5.14}$$

5.5.4 Connection between the HM-MDP and the POMDP Formulations

It can be easily seen that HM-MDP is a special case of the partially observable Markov decision processes (POMDP). Indeed, it is always possible to convert an HM-MDP into a POMDP with an augmented state space $\mathbb{S} \times \mathbb{M}$ [104, 106]. The new state space of the POMDP is the product of the original state space and mode space, containing all possible mode-state pairs. The observation space of the POMDP is the state space $\mathbb{S}$ as in the original problem the state is fully observable. The action space of the POMDP is the same with the original action space $\mathbb{A}$. The transition probability from one POMDP state (one mode-state pair for HM-MDP) to another POMDP state (another HM-MDP mode-state pair) is simply the corresponding mode transition probability $x_{mm'}$ multiplied by the corresponding state transition probability $\mathbb{P}_{ss'}^{a,m}$. This is because the state transition and mode transition are independent of each other, in spite of the fact that state transition is a function of the successor mode.

While POMDPs are superior in terms of representational power, in our non-stationary problem, HM-MDP is a more natural formulation to incorporate local information and external factors. More importantly, however, HM-MDPs require fewer parameters in model learning. In general, an HM-MDP contains $(|\mathbb{S}|^2 \cdot |\mathbb{M}| \cdot$

$|\mathbb{A}| + |\mathbb{M}|^2$) number of parameters, which is much fewer than ($|\mathbb{S}|^2 \cdot |\mathbb{M}|^2 \cdot |\mathbb{A}|$) number of parameters required for the corresponding POMDP. This simplification was shown to have significant speedup in model learning in [104, 106].

5.6 Optimal Policies for Microgrid Controller Decision Making: Exact Solution Algorithm

In this section we present an exact algorithm for solving the above formulated HM-MDP problem for an individual microgrid. From the *Optimality Principle*, the optimal value function satisfies the following backward induction recursion called the Bellman equation [107]:

$$V_t^*(s) = \max_a \left\{ R(s,a) + \gamma \sum_{s' \in \mathbb{S}} P_{ss'}^a V_{t+1}^*(s') \right\} \tag{5.15}$$

where V_t^* is the optimal value with t steps to go before termination. A difficulty in evaluating the Bellman equation for the HM-MDP model (the counterpart of (5.15) for HM-MDP) is that we cannot perform summation over the continuous joint space $\mathbb{S} \times \mathbb{B}$. However, given any pair (s, μ), since the action and state sets are finite, from (5.10) we know that there are only a finite number of possible successor pairs $(s', \mu_{s'}^a)$. Thus the Bellman equation for the HM-MDP model can be expressed as (5.16) in which the state has been replaced by the (*state*, *belief mode*) pair compared to (5.15). We can further rewrite (5.16) as (5.17) by using (5.18). Similarly, the transition probability $Prob(s'|s, \mu, a)$ and the immediate reward $R(s, \mu, a)$ can be rewritten, respectively, as (5.19) and (5.20).

$$V_t^*(s,\mu) = \max_a \left\{ R(s,\mu,a) + \gamma \sum_{(s^{'},\mu^{'})\in(\mathbb{S},\mathbb{B})} Prob(s^{'},\mu^{'}|s,\mu,a) V_{t+1}^*(s^{'},\mu^{'}) \right\} \tag{5.16}$$

$$= \max_a \left\{ R(s,\mu,a) + \gamma \sum_{s^{'}\in\mathbb{S}} Prob(s^{'}|s,\mu,a) V_{t+1}^*(s^{'},\mu_{s^{'}}^a) \right\} \tag{5.17}$$

$$Prob(s^{'},\mu^{'}|s,\mu,a) = \begin{cases} Prob(s^{'},\mu_{s^{'}}^a|s,\mu,a) = Prob(s^{'}|s,\mu,a) & \text{if } \mu^{'} = \mu_{s^{'}}^a \\ 0 & \text{if } \mu^{'} \neq \mu_{s^{'}}^a \end{cases} \tag{5.18}$$

$$Prob(s^{'}|s,\mu,a) = \sum_{m\in\mathbb{M}} Prob(s^{'}|s,m,a)\mu(m) = \sum_{m\in\mathbb{M}} \sum_{m^{''}\in\mathbb{M}} P_{ss^{'}}^{a,m^{''}} x_{mm^{''}} \mu(m) \tag{5.19}$$

$$R(s,\mu,a) = \sum_{m\in\mathbb{M}} R^m(s,a)\mu(m), \tag{5.20}$$

where $R^m(s,a)$ is the immediate reward when the action a is taken in state s while the environment mode is m.

It is worth pointing out that (5.17) still can not be directly solved easily because $V^*(t)$ is a function of continuous-valued belief mode. Therefore, a finite representation for the optimal value function $V^*(t)$ is required for the one step value function iteration in (5.17). In the rest of this section, we first present a finite representation for the optimal value function (5.17). Based on this finite representation, we will develop an update rule for the value iteration algorithm for the Bellman equation (5.17).

5.6.1 Finite Representation of Value Functions

The key to implementing a value iteration algorithm based on (5.17) is to construct a finite representation for the optimal value function $V^*(t)$. In general, the existence of such a finite representation is not guaranteed for arbitrary optimal value functions. However, a sufficient condition to guarantee the existence of a finite representation is that the optimal value function is *piecewise linear and convex* (PWLC) [108]. The convexity property implies that the value function is the upper surface of those linear value planes. Thus, in a PWLC optimal value function $V^*(s, \mu)$ with the state s given, each linear segment is a hyper-plane in an $|\mathbb{M}|$-space and can be represented by an $|\mathbb{M}|$-vector of coefficients. We say that a collection of sets $\{\Omega_s | s \in \mathbb{S}\}$ represents the value function if for any μ [98]:

$$V_t^*(s, \mu) = \max_{\omega_s \in \Omega_s} \sum_{m \in \mathbb{M}} \mu(m)\omega_s(m) = \max_{\omega_s \in \Omega_s} \mu \cdot \omega_s \tag{5.21}$$

where $\omega_s \in \Omega_s$, $\omega_s = [\omega_s(1), \omega_s(2), \ldots, \omega_s(|\mathbb{M}|)]$, and Ω_s represents the set of vectors or hyper-planes that comprise a PWLC value function $V^*(s, \mu)$. For each vector $\omega_s \in \Omega_s$, we define the *witness region* $W(\omega_s, \Omega_s)$ as a set of belief modes, in which ω_s dominates over other vectors, as shown below,

$$W(\omega_s, \Omega_s) = \{\mu | \mu \cdot \omega_s > \mu \cdot \tilde{\omega}_s, \forall \tilde{\omega}_s \in \Omega_s - \{\omega_s\}, \mu \in \mathbb{B}\} \tag{5.22}$$

A vector $\tilde{\omega}_s$ is said to be *dominated* if for $\forall \mu \in \mathbb{B}$, $\mu \cdot \tilde{\omega}_s \leq \max_{\omega_s \in \Omega_s} \mu \cdot \omega_s$. In other words, a dominated vector has an empty witness region. For any PWLC value function V_t^* that can be represented by a certain vector set collection $\{\Omega_s | s \in \mathbb{S}\}$, we can find infinitely many such representation vector sets by adding dominated vectors to Ω_s and all these set collections can represent V_t^*. However, among all these

sets there is a unique minimal set $\Omega^*_{s,t}$ for each state s, in which all vectors have nonempty witness regions. We use the term *parsimonious set* [98] when referring to this unique minimal subset of all vector sets representing a value function. It has been shown in [108] that any PWLC value function indeed has a unique parsimonious representation.

5.6.2 PWLC Properties of HM-MDP Optimal Value Functions

The parsimonious representations of optimal value functions in (5.17) provides us the possibility to update the parsimonious sets of optimal value functions in each iteration, instead of updating the value functions themselves. Before we implement this approach, however, we need the following theorem:

Proposition 2. *For an HM-MDP with arbitrary but finite scheduling period of T, the optimal finite horizon value function $V^*(t)$ is PWLC.*

Note that the HM-MDP model is a special case of the more general POMDP model [104] and the PWLC property for the POMDP has been proved before in [98]. However, in the following we present a detailed proof for the theorem proposed for the HM-MDP above because the proof itself provides an update rule for HM-MDP value function iteration based on finite representations.

$$V_t^*(s,\mu) = \max_a V_t^{a,*}(s,\mu), \qquad V_t^{*,a}(s,\mu) = \sum_{s'} V_t^{*,a,s'}(s,\mu), \tag{5.23}$$

$$V_t^{*,a,s'}(s,\mu) = \frac{1}{|\mathbb{S}|} R(s,\mu,a) + \gamma Prob(s'|s,\mu,a) V_{t+1}^*(s',\mu_{s'}^a) \tag{5.24}$$

$$V_t^{*,a,s'}(s,\mu) = \frac{1}{|\mathbb{S}|} \sum_m \mu(m) R^m(s,a) + \gamma \sum_{m'} \left(\sum_m x_{mm'} P_{ss'}^{a,m'} \mu(m) \right) \omega_{s',t+1}^*(\mu_{s'}^a, m')$$

$$= \sum_m \mu(m) \left[\frac{1}{|\mathbb{S}|} R^m(s,a) + \gamma \sum_{m'} x_{mm'} P_{ss'}^{a,m'} \omega_{s',t+1}^*(\mu_{s'}^a, m') \right] = \mu \cdot \omega_{s,t}^{a,s'}(\mu) \tag{5.25}$$

Proof. Inspired by the work in [98], we prove this theorem by induction and for each time step we break the value function (5.17) down into a series of related value functions as shown in (5.23) and (5.24), where $V_t^{*,a}(s,\mu)$ is the value of performing action a in state belief-mode pair (s,μ) at time t and then performing optimally thereafter. $V_t^{*,a,s'}(s,\mu)$ is the expected reward attributable to resultant state s' when action a is performed in state belief-mode pair (s,μ) and the optimal actions are performed thereafter. For time $t = T$, only immediate rewards matter as there is no future rewards. Based on (5.20) and (5.24) we have

$$V_T^{*,a,s'}(s,\mu) = \frac{1}{|\mathbb{S}|} \sum_{m \in \mathbb{M}} R^m(s,a)\mu(m) \tag{5.26}$$

It can be seen that $V_T^{*,a,s'}(s,\mu)$ is a linear function of μ and is therefore trivially convex. Based on the two propositions proposed in [98], which say that finite sum and finite union operation preserve the PWLC property, we know $V_T^{*,a}(s,\mu)$ and $V_T^*(s,\mu)$ are both PWLC. The inductive step assumes that $V_{t+1}^*(s,\mu)$ is PWLC and it has the finite representation as

$$V^*_{t+1}(s^{'}, \mu^a_{s'}) = \max_{\omega_{s',t+1} \in \Omega_{s',t+1}} \mu^a_{s'} \cdot \omega_{s',t+1} \tag{5.27}$$

If we let $\omega^*_{s',t+1}(\mu) = \underset{\omega_{s',t+1} \in \Omega_{s',t+1}}{\operatorname{argmax}} \mu \cdot \omega_{s',t+1}$, then we get

$$V^*_{t+1}(s^{'}, \mu^a_{s'}) = \mu^a_{s'} \cdot \omega^*_{s',t+1}(\mu^a_{s'}) \tag{5.28}$$

Substituting (5.10), (5.19), (5.20) and (5.28) into (5.24) and letting $\omega^*_{s',t+1}(\mu^a_{s'}, m^{'})$ be the $m^{'}$-th component of $\omega^*_{s',t+1}(\mu^a_{s'})$, we have (5.25). The representation vector of $V^{*,a,s^{'}}_t(s, \mu)$ is $\omega^{a,s^{'}}_{s,t}(\mu)$ with its m-th component given by

$$\omega^{a,s^{'}}_{s,t}(\mu, m) = \frac{1}{|\mathbb{S}|} R^m(s, a) + \gamma \sum_{m^{'}} x_{mm'} P^{a,m^{'}}_{ss'} \omega^*_{s',t+1}(\mu^a_{s'}, m^{'}) \tag{5.29}$$

Similarly, based on the same two propositions, we can show that the value function $V^{*,a,s^{'}}_t$ is PWLC given that V^*_{t+1} is PWLC. Thus the theorem has been proved. □

Note that though the updating rule presented in (5.25) is for finite horizon problem, it can be also useful in infinite horizon scenarios because it at least generates an improved approximation that is closer to the optimal value function.

5.6.3 Representation Set Iteration and Pruning

As mentioned earlier, our approach to obtain the optimal value function is to characterize the set of representation sets of the value function at each time instant. We achieve this by an iteration algorithm for these representation sets starting from the

time $t = T$. It is worth pointing out that the proof presented above actually provides an update rule to compute a representation set of $V_t^*(s, \mu)$, denoted by $\Omega_{s,t}$, based on the parsimonious representation of $V_{t+1}^*(s, \mu)$, denoted by $\Omega_{s,t+1}^*$. Based on this representation set iteration, the optimal value and optimal action of the proposed HM-MDP problem are given by

$$V_t^*(s, \mu) = \max_{\omega_{s,t} \in \Omega_{s,t}} \omega_{s,t} \cdot \mu, \quad a_t^*(s, \mu) = \arg\max_{a \in \mathbb{A}} V_t^{*,a}(s, \mu) \tag{5.30}$$

However, the representation sets $\Omega_{s,t}$ generated by this procedure are not necessarily parsimonious. This could make the search for optimal representation vector inefficient. Thus, it is necessary to follow the representation set iteration with a set pruning step for each time t to make the algorithm easy to implement. In the set pruning step, given an arbitrary representation set $\Omega_{s,t}$ for a PWLC value function $V_t^*(s, \mu)$, we attempt to prune the set $\Omega_{s,t}$ to the unique parsimonious representation $\Omega_{s,t}^*$.

Denoting by $\left\{\Omega_{s,t}^* | s \in \mathbb{S}\right\}$ the parsimonious representation sets, based on the decomposition presented in (5.24), we have

$$\Omega_{s,t}^* = purge\left(\cup_{a \in \mathbb{A}} \Omega_{s,t}^{*,a}\right), \quad \Omega_{s,t}^{*,a} = purge\left(\bigoplus \left\{\Omega_{s,t}^a | s \in \mathbb{S}\right\}\right) \tag{5.31}$$

where $\left\{\Omega_{s,t}^a | s \in \mathbb{S}\right\}$ contains all vectors $\omega_{s,t}^a(\mu)$ for all (s, μ) pair and the operator $\bigoplus$ is defined as the cross sum over the sets:

$$\bigoplus \left\{\Omega_{s,t}^a\right\} = \left\{ \omega_{s,t}^{*,a} | \omega_{s,t}^{*,a} = \sum_{s'} \omega_{s,t}^{*,a,s'}, s' \in \mathbb{S}, s \in \mathbb{S} \right\}. \tag{5.32}$$

The operation $purge(\tilde{\Omega}_s)$ is defined as the procedure to find the unique parsimonious subset Ω_s given an arbitrary representation set $\tilde{\Omega}_s$.

We give a brief description of notation to be used and then follow [109] in presenting the pruning algorithm in Algorithm 1. Vector comparisons are componentwise: i.e. $\omega_1 \geq \omega_2$ if and only if for all $m \in \mathbb{M}$, $\omega_1(m) \geq \omega_2(m)$. For any $m \in \mathbb{M}$, the e_m is the m-th unit vector. Set subtraction is defined as $\Omega_1 \backslash \Omega_2 = \{\omega \in \Omega_1 | \omega \notin \Omega_2\}$. The function *find-belief-mode*$(\omega_s', \tilde{\Omega}_s)$ finds a belief mode in the witness region of ω_s', which can be implemented by linear programming approaches. There are also several other algorithms to prune the given vector sets to their parsimonious subsets, such as batch enumeration [110], Sondik's one-pass algorithm [111] and witness algorithm [112]. However, the incremental pruning algorithm developed in [109] allows solving problems that could not be solved within reasonable time limits using other algorithms mentioned above.

5.6.4 Performance Analysis

We implement the above value iteration algorithm and compare its performance with a greedy algorithm and a random decision policy. In the greedy algorithm the microgrid controller makes the decision to maximize the immediate reward in the current time interval (selling all the excess energy to the Utility) in each time step, whereas in random decision policy the microgrid controller makes a random decision on the portion of excess energy to be sold. Simulation parameters of the microgrid is set as follows: renewable generation $E_r(t) \in [50kWh, 100kWh]$, storage capacity $E_s^{max} = 300kWh$, load demand $L(t) \in [50kWh, 150kWh]$. Based on the real measured load data in Texas [53], we defined 3 Gaussian distributions with different statistics for the microgrid loads, with relatively high probability on low, medium and high load demand values, respectively. Similarly we define 3 distributions for

Algorithm 1 $purge(\tilde{\Omega}_s)$

$\Omega_s \leftarrow \Phi$
for all $m \in \mathbb{M}$ **do**
  $\omega_s \leftarrow \underset{\omega_s \in \tilde{\Omega}_s}{\text{argmax}}\ e_m \cdot \omega_s;\ \ \Omega_s \leftarrow \Omega_s \cup \{\omega_s\}$
  $\tilde{\Omega}_s \leftarrow \tilde{\Omega}_s \backslash \{\omega_s\}$
end for
while $\tilde{\Omega}_s \neq \Phi$ **do**
  for all $\omega_s^{'} \in \tilde{\Omega}_s$ **do**
    $\mu \leftarrow$ *find-belief-mode*$(\omega_s^{'}, \tilde{\Omega}_s)$
    if $\mu = null$ **then**
      $\tilde{\Omega}_s \leftarrow \tilde{\Omega}_s \backslash \{\omega_s^{'}\}$
    else
      $\omega_s \leftarrow \underset{\omega_s^{'} \in \tilde{\Omega}_s}{\text{argmax}}\ \mu \cdot \omega_s^{'};\quad \Omega_s \leftarrow \Omega_s \cup \{\omega_s\}$
      $\tilde{\Omega}_s \leftarrow \tilde{\Omega}_s \backslash \{\omega_s\}$
    end if
  end for
end while
return Ω_s

renewable generations based on the Weibull distribution of wind speed [47] and the Beta distribution of solar radiation [19]. As a result, 9 environment modes in total are defined. Buying price $p_b(t)$ and selling price $p_s(t)$ are defined for each environment mode, ranging from 10 to 18 cents per kWh [113]. The expected accumulated reward with discount factor $\gamma = 0.9$ are calculated by averaging over 100 decision process realizations with the same initial conditions.

The performance comparison among three decision making strategies is shown in Fig. 5.6, with scheduling period length ranging from 2 to 10. We set the quantization

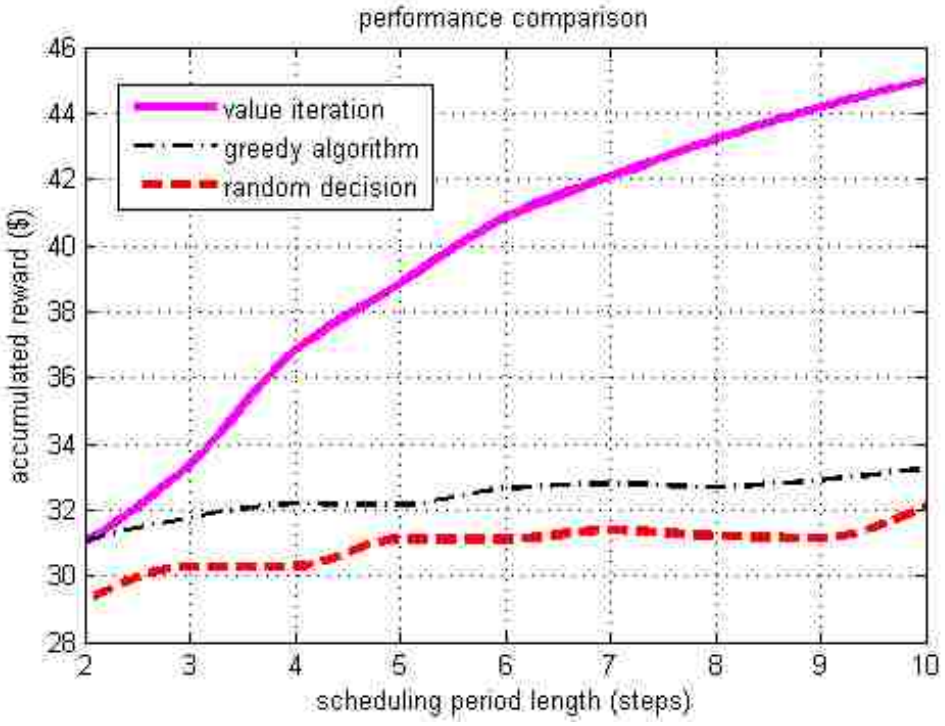

Figure 5.6: Expected accumulated reward comparison among value iteration, greedy algorithm and random decision strategy for scheduling period from 2 to 10 steps.

level for energy space as $10kWh$. Similarly, the action space $\mathbb{A} = [0, 1]$ is equally quantized with granularity of 0.1. In spite of the relatively long simulation running time, we can see that the value iteration based exact algorithm shows much better performance in accumulated rewards, no matter what the scheduling period length is.

5.7 Approximate Dynamic Programming (ADP) for Infinite Horizon Scheduling

As shown above, the value iteration algorithm leads to an optimal solution that maximizes microgrid rewards. It also enables theoretical analysis of the solution properties of the sequential decision making problem. However, there are several issues that limits its application in practice. First, it's complexity suffers from the *Curses of Dimensionality*. In our problem, the exact algorithm has exponential computational complexity in terms of the sizes of both the state space and the action space [99]. Moreover, the witness region searching becomes time inefficient when the scheduling period is long. Second, the performance of the exact algorithm depends heavily on precise transition probabilities (both state and mode transitions), which might be inaccurate as they are either estimated from the past experience or dependent on the uncertainty models which could contain modeling errors. Third, the value iteration algorithm is optimal in the specific scenario in which the sequential decisions are made in a period-by-period manner. This formulation naturally fits very well to the most common day-by-day type of scheduling, in which all decisions for the day are made at the beginning of the day without considering the information of the next day. But in more general scenarios, adaptive incorporation of new information might be required. An alternative approach that may help resolving this issue is the *receding horizon* scheme, in which a scheduling period with length T is allowed to move forward as a sliding window over time. In each step the same value iteration proposed above still applies. However, the receding horizon scheme actually pays a price for being adaptive since the computational complexity is even higher than the original day-by-day type of scheduling, especially when the sliding step size is small.

To take into account all three issues mentioned above: computational complexity, dependence on transition probability and being adaptive to new information, in this

section we resort to machine learning. In particular, we use approximate dynamic programming (ADP) and propose a Q-learning algorithm to solve the HM-MDP problem. Approximate dynamic programming is based on an algorithmic strategy that steps forward through time, by making approximations of both value functions and transition probabilities in conventional value iteration [99]. Different ADP algorithms have been proposed depending on which part of the Bellman equation (value function, policy function or transition probability) is approximated and what type of approximation is used (e.g., lookup table, parametric or nonparametric form). The ADP algorithm greatly reduces the computational complexity and approximate the transition probability to make the expectation calculation easier in the Bellman equation. Moreover, going forward in time makes it possible to expand the horizon to infinity, which overcomes the drawbacks of day-by-day and receding horizon type of scheduling schemes. It is worth pointing out that infinite horizon does not mean making plans for a infinite long time, but enables a decision making strategy to work without having to specify either starting or ending point. As a result, new observations can easily be incorporated as the algorithm moves forward over time [107].

5.7.1 Q-learning Algorithm for Model-free Decision Making

Among various ADP algorithms, Q-learning enables the microgrid controller to learn to act optimally in the presence of Markov dynamics by experiencing the consequences of actions without requiring them to build models first [114]. We define a Q-function as the value function $V^{*,a}(s,\mu)$ given the (state, belief mode) pair and the action, as shown in equation (5.33).

$$Q(s,\mu,a) = R(s,\mu,a) + \gamma \sum_{s^{'} \in \mathbb{S}} Prob(s^{'}|s,\mu,a) V^{*}(s^{'},\mu_{s^{'}}^{a}) \tag{5.33}$$

The objective in Q-learning is to estimate the above Q values for an optimal policy. To implement the Q-learning algorithm, a lookup table (Q-table) is constructed. The elements of the Q-table are the Q-values corresponding to all possible (state, mode, action) or (s, m, a) triples. In essence, the microgrid controller tries an action a when in a particular (state, belief mode) pair (s, μ) and evaluates its consequences in terms of the immediate reward it receives and its estimate of the value of the (state, belief mode) pair $(s^{'}, \mu^{a}_{s'})$ to which it is taken. The Q-values are updated according to equation (5.34) and the corresponding Q-table elements are updated according to the belief mode $\mu(m), m \in \mathbb{M}$. By trying all possible actions for all (state, belief mode) pairs repeatedly, Q-learning can learn what action is the best for each (state, belief mode) pair.

The experience of the microgrid controller consists of a sequence of distinct stages, or *episodes*, and in each episode (indexed by n), a learning rate of α_n is used to update a lookup table by incorporating new observations on reward. An exploration factor ϵ is used to balance the exploitation and exploration: The microgrid controller makes its decision according to the lookup table with probability $1-\epsilon$ and randomly explore other actions with probability ϵ. Algorithm (2) shows how the microgrid controller implements the Q-learning during the n-th episode.

Being model-free, Q-learning absorbs any changes in the non-stationary environment and provides the microgrid with an efficient algorithm to make optimal decisions when model information is limited. Generally, for a stationary MDP with fully observable states, the Q-learning algorithm indeed converges to the optimal policy if the sequence of episodes that forms the basis of learning includes an infinite number of episodes for each starting triple (s, μ, a) [107]. However, as with the POMDP problem, there is no guarantee of convergence for Q-learning in the HM-MDP problem without additional assumptions such as linear value function approximations [115]. As a result, Q-learning can be a suboptimal solution to our problem.

Algorithm 2 Q-learning

observes the current state s and estimate the belief mode pair μ

based on the pair (s, μ), the agent selects and performs an action a

observes the subsequent state $s^{'}$ and resultant estimated belief mode $\mu^{a}_{s'}$

receives an immediate reward $R(s, \mu, a)$

adjusts its old Q_{n-1} values using a learning factor α_n, according to (5.34)

$$Q_n(s,\mu,a) = \left\{ \begin{array}{ll} \left(\begin{array}{l} (1-\alpha_n)Q_{n-1}(s,\mu,a)+ \\ \alpha_n \left[R(s,\mu,a) + \gamma \max\limits_{a'} Q_{n-1}(s^{'}, \mu^{a}_{s'}, a^{'}) \right] \end{array} \right) & \text{if } \left(\begin{array}{l} (s,\mu,a) = \\ (s_n,\mu_n,a_n) \end{array} \right) \\ Q_{n-1}(s,\mu,a) & \text{otherwise} \quad , \end{array} \right\} \tag{5.34}$$

5.7.2 Performance Analysis

With the same simulation setup used in section Section IV.D, we present the performance of the Q-learning algorithm compared to the exact value iteration, greedy and the random decision algorithms in Fig. 5.7. As a balance between the convergence rate and enough weights of history data, the learning rates are set as $\alpha_n = 0.3,\ \forall n$. Before the comparison, a training period of 10^4 time steps is applied for the Q-learning algorithm with the exploration factor $\epsilon_n = 0.1,\ \forall n$ to balance the exploitation and exploration. Because of the high computational complexity of the exact algorithm, the maximum scheduling period length is 10. Though not as good as the exact algorithm, the Q-learning algorithm shows better performance than both the greedy and random algorithms.

We further investigate the performances (without the exact VI algorithm due to its high computational complexity) with longer scheduling periods (up to 20 steps),

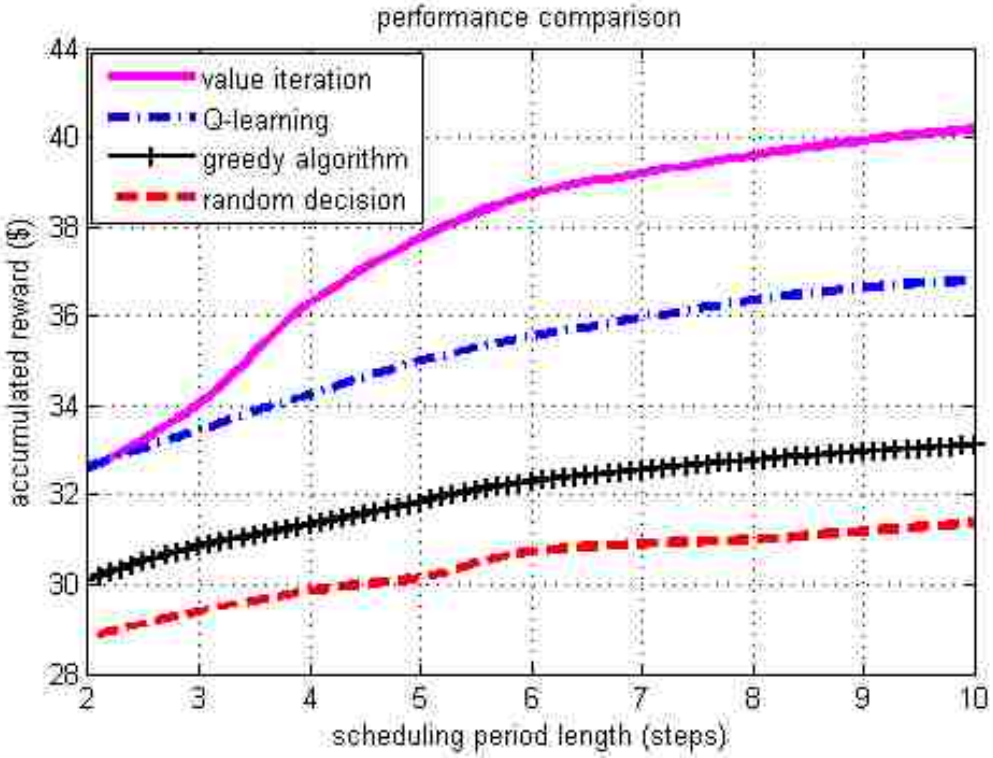

Figure 5.7: Expected accumulated reward comparison among value iteration, Q-learning, greedy and random decision algorithms for scheduling period from 2 to 10 steps.

as well as the influence of the length of the training period in Fig. 5.8. It can be seen from Fig. 8(a) that, the Q-learning algorithm shows better performance compared to the greedy algorithm and the random decisions. Moreover, it can be observed that, because of the discount factor, the accumulated reward gradually saturates as the the length of the scheduling period increases. Fig. 8(b) shows that, for the same initial parameters and a scheduling period of $T = 10$, the accumulated reward increases as the length of the training period increases from 10 to 10^4.

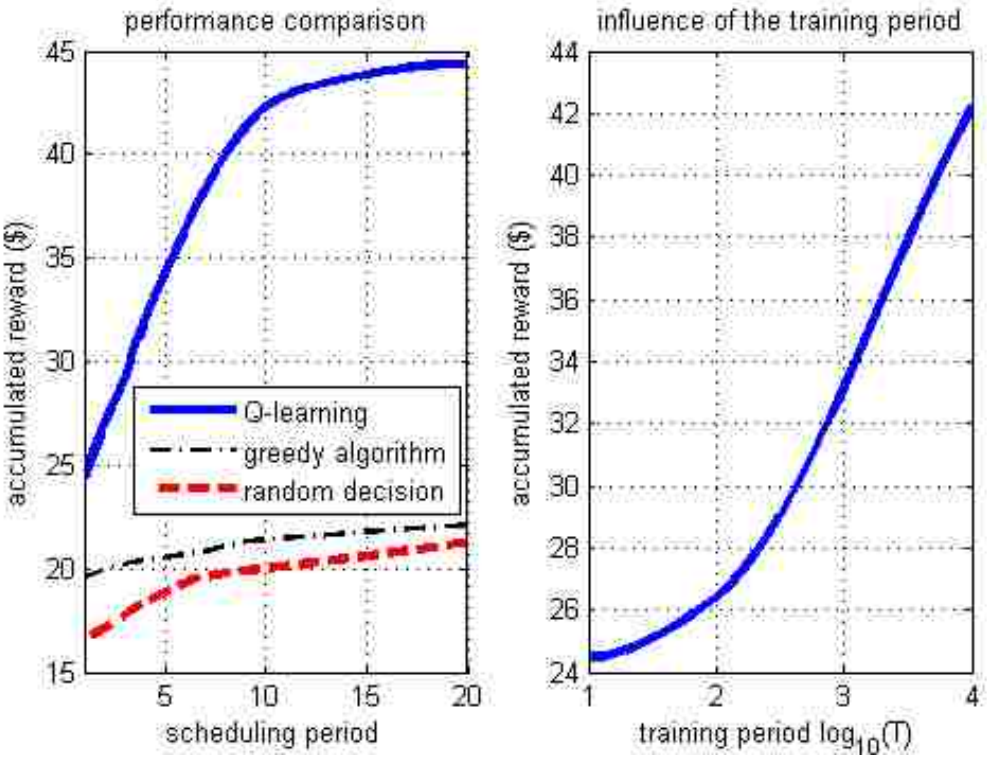

Figure 5.8: Left plot: expected return comparison among Q-learning, greedy and random decision algorithms. Right plot: the accumulated reward ($T = 10$) increases as the training period for the Q-learning increases from 10 to 10^4.

5.8 Distributed Optimal Decision Making: An Auctioning Game Design

Based on the optimal sequential decisions of the microgrid controller obtained by solving the HM-MDP model as discussed in the previous section, in this section we focus on the decision scheme design for distributed customers (smart-homes). When the microgrid controller decides to sell part of the total excess energy of the entire microgrid, this distributed decision scheme is especially important to decide how many excess energy units each smart-home contributes to the total amount of energy to be sold, considering the fact that smart-homes are all self-oriented. Several

important issues need to be addressed about the distributed decision scheme design:

1. First, the optimal distributed scheme needs to be robust to adding/removing smart-homes. This is because the number of smart-homes within a microgrid could be large and the status of smart-homes (buying-mode or selling-mode) also vary over time.

2. Second, the optimal distributed scheme needs to allow all participating selling-mode smart-homes to specify how eager they are to sell their excess energy units. Because usually not all excess energy units can be sold, a fair and efficient distributed decision scheme needs to make sure that the excess energy units to be sold are those units that the selling-mode smart-homes are highly eager to sell. Hence, to quantitatively describe the eagerness of selling-mode smart-homes, we need to define a metric of eagerness. It is worth pointing out that not only could the eagerness-metrics be different among different selling-mode smart-homes, even for the same selling-mode smart-home, the eagerness-metric might vary as the number of remaining excess energy units changes. Therefore, according to different eagerness-metrics of distributed smart-homes, the microgrid controller needs to guarantee that the energy units sold always correspond to high eagerness metrics.

3. Third, the optimal distributed scheme needs to be robust against collusive smart-homes. This is because that individual selling-mode smart-homes are all self-oriented and interested in maximizing their own benefits. Thus, selling-mode smart-homes might not necessarily telling their true eagerness metrics and they might tell the untrue values if doing so results in higher benefits.

Denote by E_l^t the total amount of excess energy (assuming E_l^t is an integer multiple of the basic energy units) that the microgrid will sell to the outside grid. Recall that E_l^t is obtained from the optimal solution of the centralized decision problem, in

which the microgrid controller decides how much electric energy need to buy or sold by the microgrid in order to maximize the accumulated reward of the entire microgrid. Since this E_l^t number of energy units comes from possibly different selling-mode smart-homes, we can define the number of *trading opportunities* to be E_l^t, where each trading opportunity corresponds to the trading of a single basic energy unit. The eagerness metric of a selling-mode smart-home associated to each trading opportunity is defined as the valuation (measured in money unit) of the trading opportunity that the smart-home has. The valuation that the smart-home has associated to an individual trading opportunity is defined as how much the smart-home expects to get from selling its excess energy unit. The valuations of the E_l^t trading opportunities are private information of smart-homes and are usually determined by factors such as energy storage, power consumptions and so on. For example, when a smart-home needs to sell its excess energy units more urgently, it will associate higher values to these trading opportunities. Hence, the eagerness-metrics it associates to these trading opportunities are also higher. With the eagerness-metric defined above, the original distributed decision making problem is equivalent to an optimal allocation problem, in which E_l^t number of trading opportunities are to be allocated among selling-mode smart-homes.

5.8.1 Vickrey Auction based Distributed Allocation Scheme

Considering all the desired properties required by the distributed decision making scheme, we propose a Vickrey auction based allocation scheme for distributed smart-homes in the microgrid, as shown in Fig. 5.2. Assume that out of the total K number of smart-homes, there are $\hat{K}_t$ number of selling-mode smart-homes in time interval t participating in the Vickrey auction competing for E_l^t number of trading opportunities. Note that in a one-shot auction in each time interval, selling excess energy always increases the smart-home's immediate reward. Thus, every smart-home wants

to sell as much excess energy as possible for its own benefit. However, from the microgrid controller's perspective, to maximize the total accumulated reward in a long run, E_l^t number of energy units must be sold to the grid at time t (this is what the first step solution determines). Since E_l^t is no greater than the total excess energy $E_x(t)$ among all smart-homes in the microgrid, only part of the excess energy units can be sold. Selling-mode smart-homes compete for the E_l^t number of trading opportunities by telling that how much money (the bids) they are willing to pay for each of the trading opportunities. Selling-mode smart-homes determine the bids based on their own valuation associated to each trading opportunity. These bids are not necessarily equal to their valuations. Thus, selling-mode smart-homes need to take into account the payments they need to make for the trading opportunities and the profits they may have by selling their excess energy units. Here the profits of a smart-home equal to the difference between the total valuations associated to all trading opportunities it obtains and the total payments it makes.

In the Vickrey auction, the k-th ($k = 1, 2, \ldots, \hat{K}_t$) selling-mode smart-home submits E_l^t number of bids $b_{t,k}^n$'s ($n = 1, 2, \ldots, E_l^t$) to indicate how much it is willing to pay for each additional trading opportunity in time interval t. Thus, bid $b_{t,k}^n$ is the amount of money the selling-mode smart-home k is willing to pay for its n-th trading opportunity. Let $\mathbf{b}_{t,k} = (b_{t,k}^1, b_{t,k}^2, \ldots, b_{t,k}^{E_l^t})$ denote the E_l^t dimensional *bid vector* with nonnegative elements of selling-mode smart-home k at time interval t. We assume that the components in the bid vector is always non-increasing in index and denote by $\mathbf{B}$ the bid vector space. $\mathbf{B}$ is a subspace of the E_l^t dimensional real vector space $\mathbb{R}_+^{E_l^t}$, which contains all E_l^t dimensional real vectors with nonnegative components. Mathematically, we have

$$\mathbf{B} := \{\mathbf{b}_{t,k} \in \mathbb{R}_+^{E_l^t} | b_{t,k}^1 \geq b_{t,k}^2 \geq \cdots \geq b_{t,k}^{E_l^t}, \forall k = 1, 2, \ldots, \hat{K}_t\} \tag{5.35}$$

Note that, in practice restricting bid vectors to have non-increasing components makes sense. This is because the selling-mode smart-home's valuations attached to individual trading opportunities is non-increasing as the smart-home gets more and more trading opportunities. For example, if a selling-mode smart-home has no trading opportunity, it needs to sell its excess energy the most urgently and its valuation for its first trading opportunity is the highest. As it sells out more and more excess energy, its storage facility gets released gradually and the marginal valuation (marginal eagerness-metric) is thus non-increasing. If a selling-mode smart-home k is only interested in selling $e_{t,k}$ ($e_{t,k} \leq E_l^t$) number of excess energy units in the auction at time t, then the last $E_l^t - e_{t,k}$ elements of its bid vector are all zeros.

A total of $\hat{K}_t \times E_l^t$ bids b_k^n's ($k = 1, 2, \ldots, \hat{K}_t; n = 1, 2, \ldots, E_l^t$) are placed for the action at time t, and the E_l^t number of trading opportunities are assigned to the E_l^t highest of these bids, which are deemed *winning bids*. Ties are broken by choosing with equal probability among all tying bids. The number of trading opportunities assigned to a selling-mode smart-home is equal to the number of winning bids submitted by that selling-mode smart-home. Thus if selling-mode smart-home k has $n_k \leq E_l^t$ of the highest bids, then it gets n_k units of trading opportunities in time interval t.

Denote by $\mathbf{c}^{-k}$ the E_l^t dimensional *competing bid* vector, which consists of the E_l^t highest others' bids, facing selling-mode smart-home k, so that c_1^{-k} is the highest of the other bids, c_2^{-k} is the second highest of the other bids, and so on. To win exactly n trading opportunities, selling-mode smart-home k's n-th highest bid must defeat the n-th lowest competing bid. If selling-mode smart-home k wins $n_{t,k}$ trading opportunities, then the the payment g_k it makes is the sum of n_k highest losing bids of the other smart-homes [116], which is given by

$$g_k = \sum_{n=1}^{n_k} g_k^n = \sum_{n=1}^{n_k} c_{E_l^t - n_k + n}^{-k}, \tag{5.36}$$

where g_k^n is the payment for the n-th trading opportunity.

5.8.2 Truthful Bidding Strategy for Vickrey Auction

In the auction in each time interval, all selling-mode smart-homes have their own valuations, which determine the bidding strategies, corresponding to all E_l^t number of trading opportunities. In the microgrid, selling-mode smart-homes do not know other's valuations precisely (*incomplete information*) since valuations of different smart-homes are determined by their own energy storage status (*private valuation*). Denoted by $\mathbf{v}_{t,k} = [v_{t,k}^1, v_{t,k}^2, \ldots, v_{t,k}^{E_l^t}]$ the private *valuation vector* of selling-mode smart-home k at time interval t, where $v_{t,k}^n$ represents the marginal value of obtaining the n-th trading opportunity. These marginal values are assumed to be non-increasing for similar reasons that we assumed non-increasing marginal bids so that

$$v_{t,k}^1 \geq v_{t,k}^2 \geq \cdots \geq v_{t,k}^{E_l^t}, \ k = 1, 2, \ldots, \hat{K}_t. \tag{5.37}$$

The total value to the selling-mode smart-home k of obtaining exactly $n_{t,k} \leq E_l^t$ trading opportunities is then the sum of the first $n_{t,k}$ marginal values: $\sum_{j=1}^{n_{t,k}} v_{t,k}^j$. Note that symmetry on valuations is usually assumed in Vickrey auction literature [116, 117], in which $\mathbf{v}_{t,k}$'s are independently and identically distributed (i.i.d) on the valuation set

$$\mathbf{V}_{t,k} = \{\mathbf{v}_{t,k} \in [0, \omega_t]^{E_l^t} : \forall n, v_{t,k}^n \geq v_{t,k}^{n+1}\}, \tag{5.38}$$

where ω_t is the maximum valuation for all selling-mode smart-homes. However, the i.i.d symmetric condition is too strong for our problem since the valuations of different selling-mode smart-homes could be different depending on individual energy consumption and storage information. Thus we drop the condition of identical distribution and assume more general asymmetric selling-mode smart-homes—smart-home k's valuation vector $\mathbf{v}_{t,k}$ is independently drawn from some distribution that has positive density everywhere on the set $\mathbf{V}_{t,k}$.

The Vickrey auction in each time interval actually forms a game with incomplete information, in which every selling-mode smart-home wants to maximize its own payoff. Here a smart-home's payoff equals the sum of valuations obtained from winning trading opportunities minus the total payment. To better analyze the formulated Vickrey auctioning game, we first introduce several important concepts from game theory and then propose an important proposition.

1. Bayesian Nash equilibrium: A Bayesian Nash equilibrium for a game with incomplete information is a strategy profile for each player that maximizes the expected payoff for each player given the strategies played by other players [116, 118, 119].

2. Strictly dominant strategy: A *strictly dominant strategy* is an action strategy that gives higher reward than any other strategy [118].

3. Weakly dominant strategy: A *weakly dominant strategy* is an action strategy that gives reward no lower than any other strategy [118].

With these concepts introduced above, we present an incentive compatibility proposition for Vickrey auction, along with its proof [116].

Proposition 3. *The Vickrey auction is incentive compatible, meaning truthful bidding (bidding the real valuation) maximizes each selling-mode smart-home's payoff*

and is a weakly dominant strategy for every selling-mode smart-home.

Proof. Consider selling-mode smart-home k and the competing bids $\mathbf{c}^{-k}$ facing it. Suppose that when smart-home k submits a bid vector $\mathbf{b}_{k,t} = \mathbf{v}_{k,t}$, it is assigned n_k trading opportunities. According to the Vickrey pricing rule, its payment is given by $\sum_{n=1}^{n_k} c^{-k}_{E^t_l - n_k + n}$ [116]. It is the case that for all $n \leq n_k$, $v^n_k \geq c^{-k}_{E^t_l - n_k + n}$ (where, $c^{-k}_{E^t_l - n_k + n} = g^n_k$), whereas for all $n > n_k$, $v^n_k < c^{-k}_{E^t_l - n_k + n}$ (where, $c^{-k}_{E^t_l - n_k + n} = g^n_k$). Now suppose selling-mode smart-home k were to submit a bid vector $\mathbf{b}_{k,t} \neq \mathbf{v}_{k,t}$ such that it is assigned the same number of trading opportunities as when it submitted its true value vector $\mathbf{v}_{k,t}$, then the payment it pays for these trading opportunities would be unaffected, as would its overall payoff. If selling-mode smart-home k were to submit a $\mathbf{b}_{k,t} \neq \mathbf{v}_{k,t}$ so that it is assigned a greater number of trading opportunities, say $n^{'}_k > n_k$, then the payments it would pay for the first n_k trading opportunities would be unchanged, and so would the payoff derived from these. For any trading opportunity $n > n_k$, the payment g^n_k exceeds (or at best equals) the n-th marginal value $v^n_{t,k}$, so the payoff from these $n^{'}_k - n_k$ trading opportunities would be negative (or at best zero). As a result, the overall surplus would be lower (or at best, the same) than that if it were to bid truthfully. Finally, if selling-mode smart-home k were to submit a $\mathbf{b}_{k,t} \neq \mathbf{v}_{t,k}$ such that it is assigned a smaller number of trading opportunities, say $n^{'}_k < n_k$, then the payments it would pay for the first $n^{'}_k$ ones would be unchanged and therefore so would the payoff derived from these. But the payoff from any trading opportunity $n < n_k$ was positive and is now forgone. Thus by winning fewer trading opportunities selling-mode smart-home k's overall payoff would be lower than if it were to bid truthfully. Based on the argument above, truthful bidding is a weakly dominant strategy for every selling-mode smart-home. □

As shown above, the truthful bidding strategy forms a Bayesian Nash equilibrium. Figure 5.9 shows an example of an individual smart-home in a Vickrey auction bidding for two trading opportunities. The competing bids from other selling-mode

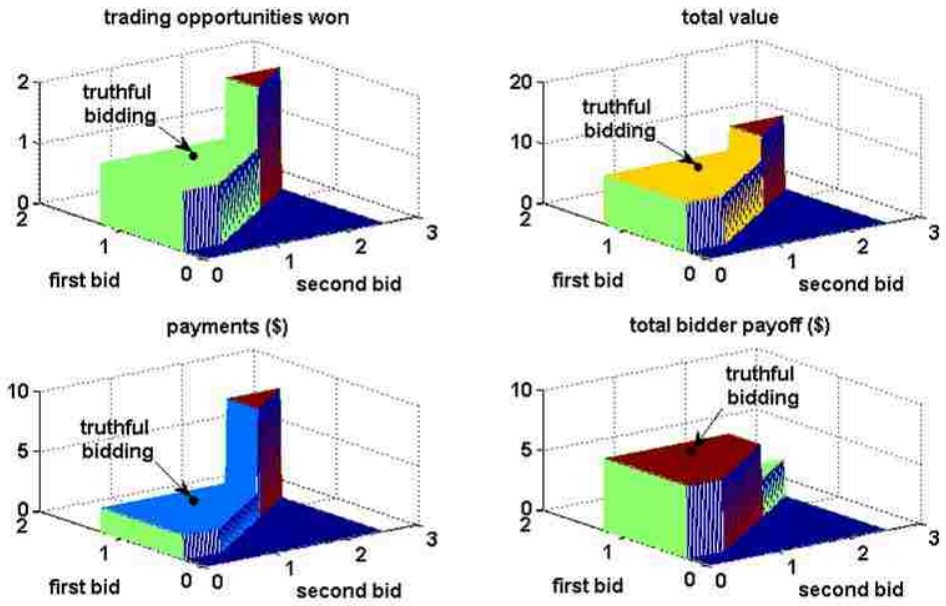

Figure 5.9: Incentive compatibility of the Vickrey auction. The normalized truthful bidding strategy point $(1, 1)$ maximizes the payoff of the individual smart-home. However, it is only weakly dominant because bidding strategies represented by other points in the same plane, within which the truthful bidding point stays, achieve the same maximum payoff.

smart-homes are assumed to be fixed. As we can see, with the bids normalized by the real valuations of the first and second trading opportunities, the truthful bidding strategy $\mathbf{b} = (1, 1)$ leads to the maximum payoff (bottom right plot). Deviation from the truthful bidding point $(1, 1)$ might increases the number of trading opportunities obtained by the smart-home (top left plot), however, the individual payoff (bottom right plot) becomes lower. Moreover, we can see that the truthful bidding strategy is only a weakly dominant strategy, because other bidding strategies represented by the points in the same plane, within which the truthful bidding point stays, achieve the same maximum payoff.

5.9 Bayesian Nash Equilibria Solution Set Structure Analysis

Though every selling-mode smart-home's payoff is maximized by the truthful bidding strategy, there is no guarantee that such Vickrey auctioning games always converge to the truthful bidding equilibrium. This is because truthful bidding is only a weakly dominant strategy and truthful bidding equilibrium is not the unique Bayesian Nash equilibrium in a Vickrey auction. Therefore detailed analysis on the entire equilibrium solution set of Vickrey auction is required.

5.9.1 The Two Types of Bayesian Nash Equilibria

Following [120], we divide the Bayesian Nash equilibria in the Vickrey auction solution set into two categories. Equilibria in the first category can be described as follows: There exists at least one selling-mode smart-home k who has at least one bid $b^n_{t,k} \in (0, \omega_t)$ with positive probability. There is a threshold $b^*_t \in (0, \omega_t)$ for all selling-mode smart-homes such that all participants bid truthfully for which they have a valuation exceeding b^*_t. Furthermore, there is a unique distinct selling-mode smart-home $\hat{k}$ who bids b^*_t on any trading opportunity for which his valuation is below the threshold. The remaining selling-mode smart-homes bid zero on any trading opportunity for which their valuation is below the threshold. Put in a more mathematical format:

$$b^n_{t,\hat{k}} = \begin{cases} v^n_{t,\hat{k}} & \text{if } v^n_{t,\hat{k}} \in [b^*_t, \omega_t] \\ b^*_t & \text{if } v^n_{t,\hat{k}} \in [0, b^*_t), \end{cases} \tag{5.39}$$

for all $n = 1, 2, \ldots, E^t_l$ and

$$b_{t,k}^n = \begin{cases} v_{t,k}^n & \text{if } v_{t,k}^n \in (b_t^*, \omega_t] \\ 0 & \text{if } v_{t,k}^n \in [0, b_t^*], \end{cases} \tag{5.40}$$

for all $k \neq \hat{k}$ and all $n = 1, 2, \ldots, E_l^t$, where ω_t is the highest valuation over all smart-homes and

$$b_t^* := \inf\{b \in (0, \omega_t) | \exists k, n \text{ s. t. } \forall \epsilon > 0, \text{Prob}\{b_{t,k}^n \in [b, b+\epsilon]\} > 0\}. \tag{5.41}$$

It can be proved that any bid strategy profile that can be described as above forms an Bayesian Nash equilibrium [120]. Conversely for any equilibrium in which certain $b_k^n \in (0, \omega_t)$ with positive probability for some selling-mode smart-home k and trading opportunity n, there is a profile of bid functions in the first category that describes the behavior of each selling-mode smart-home for almost all valuations, allowing variants (deviating behavior) on sets of measure zero of valuations. Specifically, as in reality selling-mode smart-homes usually have continuous distribution over the valuation set, there usually exists at least one selling-mode smart-home whose valuation distribution over $(0, \omega_t)$ assigns positive probability to arbitrarily small positive values. In this case, we have $b_t^* = 0$ and the first category equilibria reduce to the truthful bidding equilibrium.

For all equilibria that are not of the first type, there is zero probability of positive bids below the highest valuation ω_t. Each selling-mode smart-home k ($k = 1, 2, \ldots, \hat{K}_t$) bids at or above the highest valuation ω_t on $\hat{n}_{t,k}$ number of trading opportunities and bids zero on the remaining ones in such a manner that the total number of positive bids across all selling-mode smart-homes equals the number of trading opportunities to be sold, i.e. $\sum_{k=1}^{\hat{K}_t} \hat{n}_{t,k} = E_l^t$. The second type of Bayesian Nash equilibria reveals the possibility that the Vickrey auction might end up with a

collusive equilibrium that selling-mode smart-homes bid untruthfully and all trading opportunities are sold with zero payment.

5.9.2 Vickrey Auction Equilibrium Analysis

Vickrey auction with truthful bidding equilibrium has many good properties. For example, it is an *efficient mechanism* as it maximizes the social welfare (maximizing the sum of participants' values [116, 118, 119]). It is also incentive compatible as bidding the real values is a weakly dominant strategy for all smart-homes [116, 118, 119]. However, as mentioned above, Vickrey auction is vulnerable to collusion by selling-mode smart-homes. In the first type of Bayesian Nash equilibria, if the number of bids above the threshold is less than the number of trading opportunities for sale, then some selling-mode smart-homes will get some trading opportunities for free. In the second category of Bayesian Nash equilibria, all winning smart-homes pay zero payment for the trading opportunities they win. Generally speaking, equilibria of both categories are collusive in the sense that there are positive probabilities that smart-homes get some trading opportunities with zero payment.

The collusive equilibria jeopardize the distributed control framework in two ways: (1) The collusive equilibria fails to achieve the most important goal of the distributed decision scheme, which is to guarantee that the trading opportunities are allocated to selling-mode smart-homes who value them the highest (with highest eagerness-metric). (2) The collusive equilibria does not guarantee the profit of the auctioneer (the microgrid controller). Though in our problem, the profit of the auctioneer (the microgrid controller) is not one of the objectives to be maximized, zero payments are not desired either considering reasonable operation cost of the microgrid controller.

5.9.3 Vickrey Auction with a Reserve Price

To address the two issues raised from collusive Bayesian Nash equilibria, we further extend the Vickrey auctioning game design by introducing a reserve price. It can be proved that the Vickrey auction can be made more robust against collusive selling-mode smart-homes by introducing a positive reserve price by the microgrid controller [120, 121]. Suppose the microgrid controller sets a positive reserve price r_t for the auction in time interval t such that each selling-mode smart-home has to pay at least the reserve price for any trading opportunity obtained. Without loss of generality, bids below the reserve price, or not bidding, are identified with bidding zero. Refer to $n_t^{'}$ the number of bids at or above r_t. Then at the end of the auction, there are $\mu_t = \min\{n_t^{'}, E_l^t\}$ units are sold to the selling-mode smart-homes with the μ_t highest bids. A selling-mode smart-home who wins n_k units pays $\sum_{j=1}^{n_k} \max\{c^{-k}_{E_l^t - n_k + j}, r_t\}$. It can be shown that with a positive reserve price r_t, the Vickrey auction with more than two participants converges to a unique Bayesian Nash equilibrium, in which selling-mode smart-homes refrain from bidding on any trading opportunity for which their valuation is less than r_t and otherwise bid their valuation for each trading opportunity [120]. Introducing a reserve price not only guarantees the uniqueness of equilibrium solution of Vickrey auction, therefore making the Vickrey auction more robust to collusion by selling-mode smart-homes, but also guarantees a certain amount of benefit of the microgrid controller.

In sum, the Vickrey auction with a reserve price gives a better allocation scheme in the following aspects: (1) The Vickrey auction with a reserve price is robust to collusion by selling-mode smart-homes. (2) The Vickrey auction with a reserve price is incentive compatible, meaning assigning trading opportunities to smart-homes with highest eagerness-metrics. (3) The Vickrey auctioning game with a reserve price converges to the unique Bayesian Nash equilibrium. (4) The Vickrey auction with a reserve price guarantees a certain amount of benefit of the microgrid controller.

The only possible issue with the Vickrey auction with a reserve price is that, when the reserve price is too high, it is possible that the number of bids above the reserve price is less than E_l^t, therefore the trading opportunities assigned to selling-mode smart-homes is not enough. However, in our problem formulation, the auctioneer's profit is not one of the objectives of the distributed decision framework, thus there is no reason for the microgrid controller to set a high reserve price. In the worst case that this situation happens, repeated Vickrey auctions can be adopted and the reserve price can be adjusted until all E_l^t trading opportunities are assigned.

5.9.4 Performance Analysis

We implement the Vickrey auction (without collusion) for a microgrid model with 10 smart-homes bidding for 20 trading opportunities. The truthful valuations on the trading opportunities are within $[0, 1](\$)$. For comparison purpose, we also analyze the performance of two other auction schemes: discriminatory auction and uniform-price auction [116]. In discriminatory auction, smart-homes pay what they bid while in uniform-price auction, smart-homes pay the same highest losing bid for every trading opportunity they get. In the three different auctions, the trading opportunities, payments, payoffs of each of the 10 smart-homes, as well as the social welfare of the microgrid are compared, as shown in Fig. 5.10. Vickrey auction maximizes the social welfare of 18.87($), compared with 17.68($) of discriminatory auction and 17.87($). It is worth pointing out that in the truthful bidding equilibrium of the Vickrey auction reveals another good property in the bidding behavior of individual smart-homes, which is *Individual Rationality*, meaning the payoff function is always non-negative (as shown in the bottom plot).

In Fig. 5.11, we investigate the influence of the reserve price on the profit of the auctioneer in a one shot Vickrey auction with different time interval sizes within the

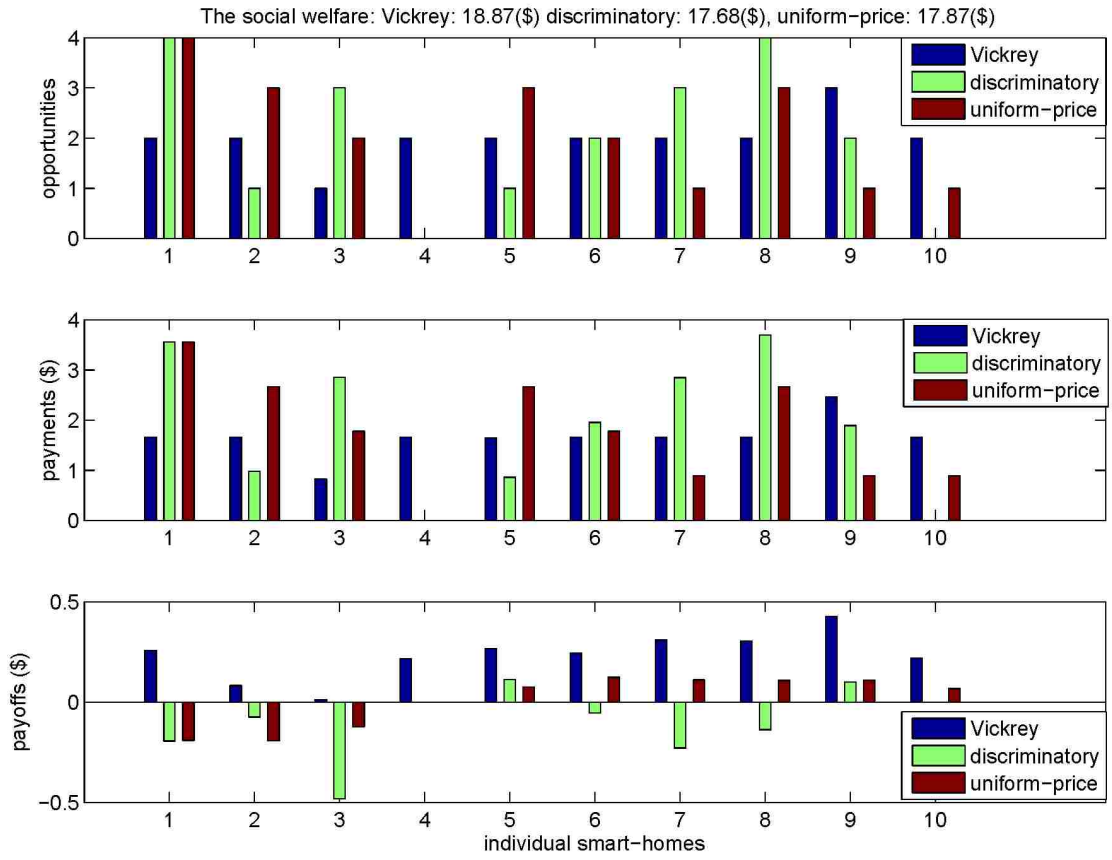

Figure 5.10: The truthful bidding equilibrium of the Vickrey auction maximizes the social welfare of the entire microgrid, while keeping the individual rationality of smart-homes.

processing block. As the reserve price (normalized by the highest value) increases from 0 to 1, after certain point, the number of trading opportunities that can be successfully allocated to smart-homes decreases from 20 to 0, which corresponds to the extreme case with reserve price higher than the highest possible value.

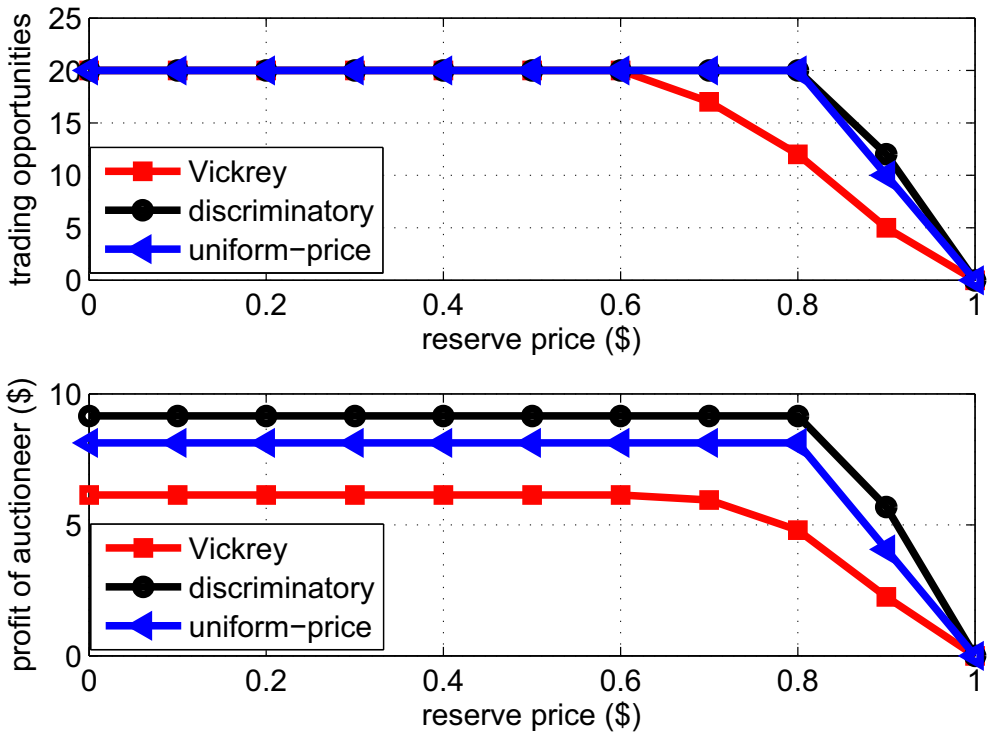

Figure 5.11: As the reserve price (normalized by the highest value) increases from 0 to 1, after certain point, the number of trading opportunities that can be successfully allocated to smart-homes decreases from 20 to 0, which corresponds to the extreme case with reserve price higher than the highest possible value.

5.10 Conclusion

In this chapter, we developed a hierarchical interactive architecture the Utility and the distributed smart-homes in a smart grid while ensuring grid-stability and Quality-of-Service (QoS). With an abstract model consisting of one controller and multiple smart-homes developed, we formulated a two-step decision framework for the real-time scheduling. The two-step decision framework consisted of (1) centralized controller sequential decisions and (2) distributed smart-home decisions. We developed a hidden mode Markov decision process (HM-MDP) model for customer real-time deci-

sion making. We first proposed a value iteration (VI) based exact solution algorithm, with the Baum-Welch and the incremental pruning (IP) algorithms adopted to learn the non-stationary dynamics and to iterate the representation sets, respectively. We further discussed the Q-learning based approximate dynamic programming (ADP) algorithm with relatively low computational complexity. Compared to greedy or random decision strategies, the Q-learning algorithm offered much more flexibility and adaptiveness with relatively good performance.

With the solution algorithm design for the HM-MDP model well developed, we then focused on the Vickrey auction design for distributed smart-homes. The solution set of the Vickrey auctioning game was divided into two categories and detailed analysis on the Bayesian Nash equilibria were presented, which showed that the truthful bidding strategy was a weakly dominant Bayesian Nash equilibrium. To overcome the vulnerability of the Vickrey auction against collusion by selling-mode smart-homes, the developed Vickrey auction was extended by introducing a reserve price, which guaranteed robustness of the auction and the convergence of the auctioning game to the unique truthful bidding equilibrium.

Chapter 6

Summary of the Dissertation and Research Directions

In this dissertation, we have developed a hierarchical interactive architecture for future smart grids. In the followings, we summarize the main aspects and contributions of this dissertation. We also propose possible research directions that can be addressed in the near future.

6.1 Summary of the Dissertation

In Chapter 2, we proposed two types of approaches to model the uncertainty in customer load demand. The first approach was based on a first order non-stationary Markov chain. A maximum likelihood estimator (MLE) was derived to estimate the time variant transition matrix of the Markov chain. The second approach was based on time series analysis techniques. We presented linear prediction models such as standard autoregressive (AR) process and time varying autoregressive (TVAR) process, according to different assumptions on the stationarity of customer load profile:

piecewise stationarity, local stationarity and cyclo-stationarity. Prediction performances of different models were analyzed and compared, advantages and disadvantages were discussed.

In Chapter 3, we designed a DR scheduling scheme based on the Utility cost minimization with different customer clustering sizes. A convex optimization problem was formulated and the optimal demand response profile was in the form of a two-dimensional *water-filling solution* either with flat water levels or different water levels for different customers. Price of Anarchy (PoA) analysis was presented to balance both the centralized and distributed competing objectives.

In Chapter 4, an optimal stochastic tracking scheme was proposed in an interactive smart grid infrastructure. Optimal stochastic control schemes for the active power control (primary frequency control) were designed, in the presence of uncertainties arising from customer load demands and distributed renewable generations, to stabilize frequency and maintain a balance between generation and consumption within the distributed synchronous area. We proposed two stochastic tracking schemes based on the state-space representation of a synchronous generator: (1) reference dynamics-based tracking and (2) reference statistics-based tracking. We further extended the proposed optimal controllers by considering the realistic scenario of asynchronous load demand signals from different households. To compensate for different delays seen by different household signals, a Kalman filter (KF) based prediction scheme was proposed to generate the correct reference signal and we showed that the centralized reference prediction could equivalently be implemented distributively. Simulation results were presented to show the performances of the proposed prediction and tracking schemes.

In Chapter 5, we developed a hierarchical interactive architecture the Utility and the distributed smart-homes in a smart grid while ensuring grid-stability and Quality-of-Service (QoS). With an abstract model consisting of one controller and

multiple smart-homes developed, we formulated a two-step decision framework for the real-time scheduling. The two-step decision framework consisted of (1) centralized controller sequential decisions and (2) distributed smart-home decisions. We developed a hidden mode Markov decision process (HM-MDP) model for customer real-time decision making. We first proposed a value iteration (VI) based exact solution algorithm, with the Baum-Welch and the incremental pruning (IP) algorithms adopted to learn the non-stationary dynamics and to iterate the representation sets, respectively. We further discussed the Q-learning based approximate dynamic programming (ADP) algorithm with relatively low computational complexity. Compared to greedy or random decision strategies, the Q-learning algorithm offered much more flexibility and adaptiveness with relatively good performance. With the solution algorithm design for the HM-MDP model well developed, we then focused on the Vickrey auction design for distributed smart-homes. The solution set of the Vickrey auctioning game was divided into two categories and detailed analysis on the Bayesian Nash equilibria were presented, which showed that the truthful bidding strategy was a weakly dominant Bayesian Nash equilibrium. To overcome the vulnerability of the Vickrey auction against collusion by selling-mode smart-homes, the developed Vickrey auction was extended by introducing a reserve price, which guaranteed robustness of the auction and the convergence of the auctioning game to the unique truthful bidding equilibrium.

6.2 Future Research Directions

The work that is presented in this dissertation can be extended along several directions, focusing on either frequency control, load prediction or smart-home decision making.

6.2.1 Particle Filtering

In Chapter 4, we did the linearization of the nonlinear dynamics of the synchronous generator at the operation point. One extension is to consider the nonlinear dynamics of the generator and design both optimal nonlinear controller and system state estimator. Particle filtering technique is one option for the nonlinear tracking control with hidden system state.

6.2.2 Cyclo-stationarity in Load Demand Prediction

It is worth pointing out that, in Chapter 2, the prediction under the cyclostationarity assumption gradually becomes off from the real measured data when the prediction period becomes too long, i.e. over months or seasons. This is because customer power consumption pattern do change from season to season. Note that verifying the cyclostationarity of a data sequence in a meaningful way is not an easy topic and out of the scope of this work, since huge amount of other information is required to go with it. Interesting readers are referred to [59] and related references there. Hence even though the TVAR model increases the prediction efficiency, updates on the model coefficients are still necessary periodically (say, monthly) in practice.

6.2.3 Decentralized Partially Observable Markov Decision Process (Dec-POMDP) in Smart-home Decision Making

In designing the distributed smart-home decision making schemes, another option is to model the decision process as a Dec-POMDP, instead of applying game theoretic approach. There can be different problem formulations depending on how we define

the objective function, observation function and action function. Decentralized control of cooperative systems captures the operation of a group of decision-makers that share a single global objective. The difficulty in solving optimally such problems arises when the agents lack full observability of the global state of the system when they operate. The general problem has been shown to be NEXP-complete [122].

Appendices

A Proof of the Separation Principle

Appendix A

Proof of the Separate Design of Control and Reference Prediction

The conventional generation facilities system dynamics can be written as

$$x(i+1) = Ax(i) + Bu(i) + w_1(i) \tag{A.1}$$

$$y(i) = Cx(i) + w_2(i) \tag{A.2}$$

$$z(i) = Cx(i) \tag{A.3}$$

where $z(i)$ is the active power generated by the conventional generation facilities and $y(i)$ is the noisy observation. Similarly, we may define the reference system dynamics, which is an augmented system incorporating all customers, as

$$x_r(i+1) = A_r x_r(i) + w_r(i) \tag{A.4}$$

$$y_r(i) = C_r x_r(i) + v_r(i) \tag{A.5}$$

$$z_r(i) = C_r x_r(i) \tag{A.6}$$

where $w_r(i)$ and $v(i)$ are process and measurement noises, both of which are assumed to be white Gaussian noises.

The objective function we want to minimize is given as follows, which is a conditional mean given the system output observation $y(i)$ and the delayed reference signal $y_r(i-d)$.

$$U = E\{\sum_{i=i_0}^{i_1} \left(z(i) - z_r(i)\right)^T R_1 \left(z(i) - z_r(i)\right) + u^T(i) R_2 u(i) | y(i), y_r(i-d)\} \tag{A.7}$$

where the control accuracy matrix R_1 is positive semi-definite and the control effort matrix R_2 is positive definite. Thus, to solve this stochastic tracking control problem, a prediction\estimation scheme is required to recover the non-delayed reference signal before any control can be implemented. This "prediction before control" structure is pretty natural and intuitive. We can show that the original objective function can be equivalently transformed into this structure.

A.1 Direct Proof of Separate Design of Control and Reference Prediction

Given the objective function (A.7), we define the Hamiltonian as

$$
\begin{aligned}
&H\left(x^*(i), u^*(i), \lambda^*(i+1)\right) = \\
&E\{\left(z^*(i) - z_r(i)\right)^T R_1 \left(z^*(i) - z_r(i)\right) + (u^*)^T(i) R_2 u^*(i) \\
&+ \left(\lambda^*(i+1)\right)^T [Ax^*(i) + Bu^*(i) + w_1(i)] | y(i), y_r(i-d)\}
\end{aligned} \tag{A.8}
$$

From the fundamental theorem of the calculus of variations, we know that for extremization of functionals, the first variation must be equal to zero. Thus we can apply the Euler-Lagrange (EL) equation to the Hamiltonian H with respect to the variables $x(i)$, $u(i)$ and $\lambda(i)$. Thus, we get

$$
\lambda^*(i) = \frac{\partial H\left(x^*(i), u^*(i), \lambda^*(i+1)\right)}{\partial x^*(i)} \tag{A.9}
$$

$$
0 = \frac{\partial H\left(x^*(i), u^*(i), \lambda^*(i+1)\right)}{\partial u^*(i)} \tag{A.10}
$$

$$
E\{x^*(i+1)|y(i+1)\} = \frac{\partial H\left(x^*(i), u^*(i), \lambda^*(i+1)\right)}{\partial \lambda^*(i)} \tag{A.11}
$$

Substituting $z(i) = Cx(i)$ into the equations above and exchanging the order of taking derivatives and expectations, we have the equations follows,

$$
\begin{aligned}
&\lambda^*(i) = 2C^T R_1 C E\{x^*(i)|y(i)\} \\
&- 2C^T R_1 E\{z_r(i)|y_r(i-d)\} + A^T \lambda^*(i+1)
\end{aligned} \tag{A.12}
$$

$$
0 = 2R_2 u^*(i) + B^T \lambda^*(i+1) \tag{A.13}
$$

$$
E\{x^*(i+1)|y(i+1)\} = AE\{x^*(i)|y(i)\} + Bu^*(i) \tag{A.14}
$$

Note that in derivation above, we drop some condition terms in the conditional expectation, considering the controlled system and the reference system are independent with each other, which means $E\{x^*(i)|y(i), y_r(i-d)\} = E\{x^*(i)|y(i)\}$ and $E\{z_r(i)|y(i), y_r(i-d)\} = E\{z_r(i)|y_r(i-d)\}$. The noise term is also dropped because of the white Gaussian noise assumption.

Combining (A.13) and (A.14) we have

$$\begin{aligned} &E\{x^*(i+1)|y(i+1)\} = \\ &AE\{x^*(i)|y(i)\} - \frac{1}{2}BR_2^{-1}B^T\lambda^*(i+1) \end{aligned} \tag{A.15}$$

From (A.13) we have the open loop optimal control law as

$$u^*(i) = -\frac{1}{2}R_2^{-1}B^T\lambda^*(i) \tag{A.16}$$

and optimal costate $\lambda^*(i)$ and optimal state $x^*(i)$ can be obtained by iteratively solving the difference equations (A.12) and (A.15).

In order to obtain the closed-loop configuration, we need to try to express the costate $\lambda^*(i)$ in the optimal control (A.16) in terms of the state estimate $E\{x^*(i)|y(i)\}$. The final condition

$$\lambda(i_1) = R_1E\{x(i_1)|y(i_1)\}) \tag{A.17}$$

prompts us to express

$$\lambda(i) = P(i)E\{x(i)|y(i)\}), \tag{A.18}$$

where $P(i)$ is yet to be determined. This linear transformation is called the Ricatti transformation.

By doing some straight forward derivation, we have

$$\begin{aligned} &P(i)E\{x^*(i)|y(i)\} = \\ &R_1E\{x^*|y(i)\} - 2C^TR_1E\{z_r(i)|y_r(i-d)\} \\ &+ A^TP(i+1)[I + \frac{1}{2}BR_2^{-1}B^TP(i+1)]^{-1} \\ &\cdot AE\{x^*(i)|y(i)\} \end{aligned} \tag{A.19}$$

Given the boundary condition, we can solve this difference equation for $P(i)$ iteratively. Thus, the closed-loop optimal control law is given by

$$\begin{aligned} u^*(i) = -\frac{1}{2}R_2^{-1}B^TA^{-T}[(P(i) \\ - 2C^TR_1C)E\{x^*|y^*\} - 2C^TR_1C_rE\{x_r(i)|y_r(i-d)\}] \end{aligned} \tag{A.20}$$

Now we can see that the optimal control input to minimize the objective function (A.7) can be find based on the conditional means of the system state and the reference signal, which can be obtained by a minimum mean square error estimator. Since the system and reference dynamics are all linear systems, then the Kalman filter is the best linear MMSE. To sum up, in the original optimal tracking control scheme design with asynchronous reference signal, the predictor design for the reference prediction, observer design for the state estimation and controller design for output tracking control can be implemented separately. This separation principle of reference prediction, state observation and controller design is overall optimal for the original objective function (A.7).

A.2 Backward Proof of the Separate Design in Control and Reference Prediction

Rewrite the objective function as $U = \sum_{i=i_0}^{i_1} J(i)$ and $J(i) = E\{(z-z_r)^T R_1(z-z_r)+ u^T(i)R_2u(i)|y(i), y_r(i-d)\}$. Substituting $z(i) = Cx(i)$ and $z_r(i) = C_r x_r(i)$, then we have

$$\begin{aligned} J(i) &= E\{(z(i) - z_r(i))^T R_1(z(i) - z_r(i)) + u^T(i)R_2u(i)|y(i), y_r(i-d)\} \\ &= E\{(z(i) - C_r x_r(i))^T R_1(z(i) - C_r x_r(i)) + u^T(i)R_2u(i)|y(i), y_r(i-d)\} \end{aligned} \tag{A.21}$$

Denote by $\hat{x}_r(i)$ the MMSE estimate of $x_r(i)$ given the observation of the delayed reference signal $y_r(i-d)$, denote by $\tilde{x}_r(i)$ the estimation error. Then we have

$$\begin{aligned} J(i) = &E\{(z(i) - C_r\hat{x}_r(i))^T R_1(z(i) - C_r\hat{x}_r(i))|y(i), y_r(i-d)\} \\ &- E\{(C_r\tilde{x}_r(i))^T R_1(z(i) - C_r\hat{x}_r(i))|y(i), y_r(i-d)\} \\ &- E\{(z(i) - C_r\hat{x}_r(i))^T R_1(C_r\tilde{x}_r(i))|y(i), y_r(i-d)\} \\ &+ E\{(C_r\tilde{x}_r(i))^T R_1(C_r\tilde{x}_r(i))|y(i), y_r(i-d)\} \\ &+ E\{u^T(i)R_2u(i)|y(i), y_r(i-d)\} \end{aligned} \tag{A.22}$$

Considering that the controlled system and the reference system are independent with each other, we can drop some condition terms in the above expectations. Note that the second and third terms

$$E\{(C_r\tilde{x}_r(i))^T R_1(z(i) - C_r\hat{x}_r(i))|y(i), y_r(i-d)\} \tag{A.23}$$

and

$$E\{(z(i) - C_r\hat{x}_r(i))^T R_1(C_r\tilde{x}_r(i))|y(i), y_r(i-d)\} \tag{A.24}$$

are both zeros. This can be shown by taking into account the zero mean property of the estimation error and the orthogonality principle of the MMSE estimator. We take the second term as an example.

$$E\{(C_r\tilde{x}_r(i))^T R_1(z(i) - C_r\hat{x}_r(i))|y(i), y_r(i-d)\} \tag{A.25}$$

$$= E\{(\tilde{x}_r(i))^T (C_r)^T R_1 z(i)|y(i), y_r(i-d)\} -$$

$$E\{(\tilde{x}_r(i))^T (C_r)^T R_1 C_r\hat{x}_r(i)|y(i), y_r(i-d)\} \tag{A.26}$$

$$= (E\{\tilde{x}_r(i)\})^T (C_r)^T R_1 z(i) - \mathrm{tr}\{(C_r)^T R_1 C_r E\{\hat{x}_r(i)(\tilde{x}_r(i))^(T)\}|y(i), y_r(i-d)\} \tag{A.27}$$

$$= 0 \tag{A.28}$$

The forth expectation term $E\{(C_r\tilde{x}_r(i))^T R_1(C_r\tilde{x}_r(i))|y(i), y_r(i-d)\}$ can be written as $\mathrm{tr}\{(C_r)^T R_1(C_r)V\}$ where $V = E\{\tilde{x}_r(i)(\tilde{x}_r(i))^T|y(i), y_r(i-d)\}$ is the conditional covariance matrix of the reconstruction error $\tilde{x}_r(i)$.

Thus the original objective function (A.7) can be equivalently transformed to the following form, with the MMSE estimator applied.

$$\begin{aligned} U = \sum_{i=i_0}^{i_1} E\{[(z(i) - C_r\hat{x}_r(i))^T R_1(z(i) - C_r\hat{x}_r(i)) \\ + u^T(i)R_2u(i)]|y(i), y_r(i-d)\} + \mathrm{tr}\{\sum_{i=i_0}^{i_1}[(C_r)^T R_1(C_r)V]\} \end{aligned} \tag{A.29}$$

It can be seen that the last term in this expression is independent of the control applied to the system. The objective function (A.29) (as well as (A.7)) can be minimized by the control input u that minimizes the first expectation term

$$E\{\sum_{i=i_0}^{i_1}\left[(z-C_r\hat{x}_r)^T R_1(z-C_r\hat{x}_r)+u^T R_2 u\right]|y(i),y_r(i-d)\} \tag{A.30}$$

which is a standard linear quadratic output tracking problem with the MMSE estimate $\hat{z}_r$ as the tracking reference signal.

Note that in the proof above we assume we have perfect observation of the state, if the system state is not directly available, an optimal observer is needed to reconstruct the state for the controller design. Similarly we can show that the separation principle also holds for the observer and controller design. To sum up, in the original optimal tracking control scheme design with asynchronous reference signal, the predictor design for the reference prediction, observer design for the state estimation and controller design for output tracking control can be implemented separately. This separation principle of reference prediction, state observation and controller design is overall optimal for the original objective function (A.7).

References

[1] E. Santacana, G. Rackliffe, L. Tang, and X. Feng, "Getting smart," *IEEE Power and Energy Magazine*, vol. 8, no. 2, pp. 41 – 48, Mar 2010.

[2] "The Smart Grid: An Introduction," The U.S. Department of Energy, Tech. Rep., Oct. 2008.

[3] J. A. Momoh, "Smart grid design for efficient and flexible power networks operation and control," in *2009 IEEE/PES Power Systems Conference and Exposition, PSCE '09.*, Mar. 2009, pp. 1 –8.

[4] S. K. Mazumder, K. Acharya, and M. Tahir, "Towards realization of a control-commnunication framework for interactive power networks," in *2008 IEEE Energy 2030 Conference*, Nov. 2008, pp. 1 – 8.

[5] J. Taneja, D. Culler, and P. Dutta, "Towards cooperative grids: Sensor/actuator networks for renewables integration," in *2010 First IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Oct. 2010, pp. 531 –536.

[6] I. P. Horacek, *Power Balance Control in Electrical Grids*. New York: Wiley Interscience, 2010.

[7] "Balancing and frequency control," North American Electric Reliability Corporation (NERC), Tech. Rep., Jan. 2011.

[8] N. Jaleeli, L. S. VanSlyck, D. N. Ewart, L. H. Fink, and A. G. Hoffmann, "Understanding automatic generation control," *IEEE Transactions on Power Systems*, vol. 7, no. 3, pp. 1106 – 1122, Aug. 1992.

[9] R. Albert, I. Albert, and G. L. Nakarado, "Structural vulnerability of the north american power grid," *Physical Review E*, vol. 69, p. 025103, Feb. 2004.

[10] P. Jarventausta, S. Repo, A. Rautiainen, and J. Partanen, "Smart grid power system control in distributed generation environment," *Annual Reviews in Control*, vol. 34, no. 2, pp. 277 – 286, 2010.

[11] W. Kempton and J. Tomic, "Vehicle-to-grid power implementation: From stabilizing the grid to supporting large-scale renewable energy," *Journal of Power Sources*, vol. 144, no. 1, pp. 280–294, 2005.

[12] C. W. Potter, A. Archambault, and K. Westrick, "Building a smarter smart grid through better renewable energy information," in *2009 IEEE/PES Power Systems Conference and Exposition, PSCE '09.*, Mar. 2009, pp. 1–5.

[13] K. J. P. Macken, K. Vanthournout, J. V. den Keybus, G. Deconinck, and R. J. M. Belmans, "Distributed control of renewable generation units with integrated active filter," *IEEE Transactions on Power Electronics*, vol. 19, no. 5, pp. 1353 – 1360, Sept. 2004.

[14] D. Li and S. K. Jayaweera, "Distributed smart-home decision-making in a hierarchical interactive smart grid architecture," *IEEE Transactions on Parallel and Distributed Systems*, Feb. 2014, accepted for publication.

[15] H. Sui, H. Wang, M.-S. Lu, and W.-J. Lee, "An ami system for the deregulated electricity markets," *IEEE Transactions on Industry Applications*, vol. 45, no. 6, pp. 2104 – 2108, Nov 2009.

[16] S. K. Jayaweera and D. Li, "Optimal control of smart-grid with asynchronous distributed renewables," in *the 4th International Conference on Intelligence and Advanced Systems (ICIAS2012)*, Perak, Malaysia, Jun. 2012.

[17] J. P. Barton and D. G. Infield, "Energy storage and its use with intermittent renewable energy," *IEEE Transactions on Energy Conversion*, vol. 19, no. 2, pp. 441 – 448, Jun. 2004.

[18] R. Karki, H. Po, and R. Billinton, "A simplified wind power generation model for reliability evaluation," *IEEE Transactions on Energy Conversion*, vol. 21, no. 2, pp. 533 – 540, Jun. 2006.

[19] Y. M. Atwa, E. F. El-Saadany, M. M. A. Salama, and R. Seethapathy, "Optimal renewable resources mix for distribution system energy loss minimization," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 360 –370, Feb. 2010.

[20] R. Yao and K. Steemers, "A method of formulating energy load profile for domestic buildings in the uk," *Energy and Buildings*, vol. 37, no. 6, pp. 663 – 671, 2005.

[21] J. V. Paatero and P. D. Lund, "A model for generating household electricity load profiles," *International Journal of Energy Research*, vol. 30, no. 5, pp. 273–290, 2006.

[22] I. F. Visconti, L. F. W. de Souza, J. M. S. C. Costa, and N. R. B. C. Sobrinho, "From power quality monitoring to transient stability analysis: Measurement-based load modeling for dynamic simulations," in *2010 14th International Conference on Harmonics and Quality of Power*, Bergamo, Italy, Sept. 2010, pp. 1 – 7.

[23] "Benefits of demand response in electricity markets and recommendations for achieving them," U.S. Department of Energy, Tech. Rep., Feb. 2006.

[24] N. Lu, D. P. Chassin, and S. E. Widergren, "Modeling uncertainties in aggregated thermostatically controlled loads using a state queueing model," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 725–733, May 2005.

[25] S. H. S. Han and K. Sezaki, "Development of an optimal vehicle-to-grid aggregator for frequency regulation," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 65–72, Apr. 2010.

[26] F. Rahimi and A. Ipakchi, "Demand response as a market resource under the smart grid paradigm," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 82 – 88, Jun. 2010.

[27] D. P. Chassin, D. J. Hammerstrom, and J. G. DeSteese, "The Pacific Northwest demand response market demonstration," *IEEE PES GM*, pp. 1 – 6, Jul. 2008.

[28] Y. G. Rebours, D. S. Kirschen, M. Trotignon, and S. Rossignol, "A survey of frequency and voltage control ancillary services - part i: Technical features," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 350 – 357, Feb. 2007.

[29] M. Albadi and E. El-Saadany, "A summary of demand response in electricity markets," *Electric Power Systems Research*, vol. 78, no. 11, pp. 1989 – 1996, 2008.

[30] D. Li, S. K. Jayaweera, O. Lavrova, and R. Jordan, "Load management for price-based demand response scheduling - a block scheduling model," in *International Conference on Renewable Energies and Power Quality (ICREPQ'11)*, Las Palmas de Gran Canaria, Spain, Apr. 2011.

[31] D. Li, S. K. Jayaweera, and A. Naseri, "Auctioning game based demand response scheduling in smart grid," in *2011 IEEE Online Conference on Green Communications (GreenCom)*, Sept. 2011, pp. 58 – 63.

[32] A. Molderink, V. Bakker, M. G. C. Bosman, J. L. Hurink, and G. J. M. Smit, "A three-step methodology to improve domestic energy efficiency," in *2010 Innovative Smart Grid Technologies (ISGT)*, Washington D.C., USA, Jan. 2010, pp. 1 – 8.

[33] H. D. Keulenaer, "Leonardo Energy Minute Lecture: Voltage and Frequency Control," 2007, http://www.leonardo-energy.org/voltage-and-frequency-control-grid.

[34] N. C. Scott, D. J. Atkinson, and J. E. Morrell, "Use of load control to regulate voltage on distribution networks with embedded generation," *IEEE Transactions on Power Systems*, vol. 17, no. 2, pp. 510 –515, May 2002.

[35] R. G. de Almeida and J. A. P. Lopes, "Participation of doubly fed induction wind generators in system frequency regulation," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 944 –950, Aug. 2007.

[36] B. V. Perumal and J. K. Chatterjee, "Voltage and frequency control of a stand alone brushless wind electric generation using generalized impedance controller," *IEEE Transactions on Energy Conversion*, vol. 23, no. 2, pp. 632 – 641, Jun. 2008.

[37] I. Tamrakar, L. B. Shilpakar, B. G. Fernandes, and R. Nilsen, "Voltage and frequency control of parallel operated synchronous generator and induction generator with statcom in micro hydro scheme," *IET Generation, Transmission Distribution*, vol. 1, no. 5, pp. 743 – 750, Sept. 2007.

[38] I. Erlich, K. Rensch, and F. Shewarega, "Impact of large wind power generation on frequency stability," in *IEEE Power Engineering Society General Meeting*, Jun. 2006, p. 8 pp.

[39] J. Kanniah, O. P. Malik, and G. S. Hope, "Excitation control of synchronous generators using adaptive regulators part i-theory and simulation results," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-103, no. 5, pp. 897 – 903, May 1984.

[40] G. Raina, O. P. Malik, and A. Kumar, "Wind oscillatory and gusting torque effect on a power system," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-21, no. 6, pp. 767 – 776, Nov. 1985.

[41] O. I. Elgerd and C. Fosha, "Optimum megawatt-frequency control of multiarea electric energy systems," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, no. 4, pp. 556 – 563, Apr. 1970.

[42] M. A. Ortega-Vazquez and D. S. Kirschen, "Optimizing the spinning reserve requirements using a cost/benefit analysis," *IEEE Transactions on Power Systems,*, vol. 22, no. 1, pp. 24 –33, Feb. 2007.

[43] M. Walsh and M. J. O'Malley, "Augmented hopfield network for unit commitment and economic dispatch," *IEEE Transactions on Power Systems*, vol. 12, no. 4, pp. 1765 –1774, Nov. 1997.

[44] D.-M. Han and J.-H. Lim, "Design and implementation of smart home energy management systems based on zigbee," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 3, pp. 1417 –1425, Aug. 2010.

[45] Y.-J. Lin, H. A. Latchman, M. Lee, and S. Katar, "A power line communication network infrastructure for the smart home," *IEEE Wireless Communications Magazine*, vol. 9, no. 6, pp. 104 – 111, Dec. 2002.

[46] D. Pudjianto, C. Ramsay, and G. Strbac, "Virtual power plant and system integration of distributed energy resources," *IET Renewable Power Generation*, vol. 1, no. 1, pp. 10 – 16, Mar. 2007.

[47] B. Safari and J. Gasore, "A statistical investigation of wind characteristics and wind energy potential based on the weibull and rayleigh models in rwanda," *Renewable Energy*, vol. 35, no. 12, pp. 2874 – 2880, 2010.

[48] S. Shao, M. Pipattanasomporn, and S. Rahman, "Challenges of phev penetration to the residential distribution network," in *IEEE Power Energy Society General Meeting, 2009. PES '09.*, Jul. 2009, pp. 1 –8.

[49] K. J. Dyke, N. Schofield, and M. Barnes, "The impact of transport electrification on electrical networks," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 12, pp. 3917 –3926, Dec. 2010.

[50] W. Tang and R. Jain, "Stochastic resource auctions for renewable energy integration," in *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2011, pp. 345–352.

[51] B. Ramachandran, S. K. Srivastava, D. Cartes, and C. Edrington, "Distributed energy resource management in a smart grid by risk based auction strategy for profit maximization," in *2010 IEEE Power and Energy Society General Meeting*, 2010, pp. 1–7.

[52] W. Saad, Z. Han, H. V. Poor, and T. Basar, "A noncooperative game for double auction-based energy trading between phevs and distribution grids," in *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2011, pp. 267–272.

[53] "Backcasted (Actual) Load Profiles - Historical," http://www.ercot.com/mktinfo/loadprofile/alp/.

[54] T. S. Rao, "The fitting of non-stationary time-series models with time-dependent parameters," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 32, no. 2, pp. 312–322, 1970.

[55] S. Degerine and S. Lambert-Lacroix, "Characterization of the partial autocorrelation function of nonstationary time series," *Journal of Multivariate Analysis*, vol. 87, no. 1, pp. 46 – 59, 2003.

[56] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*. 111 River Street, Hoboken, NJ: John Wiley & Sons, 2008.

[57] M. De-Hoon, T. V. der Hagen, H. Schoonewelle, and H. Van-Dam, "Why yulecwalker should not be used for autoregressive modeling," *Annals of Nuclear Energy*, vol. 23, no. 15, p. 1219C1228, Feb. 1996.

[58] Y. Grenier, "Time-dependent ARMA modeling of nonstationary signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, no. 4, pp. 899 – 911, Aug. 1983.

[59] W. Gardner, "Measurement of spectral correlation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, no. 5, pp. 1111 – 1123, Oct. 1986.

[60] A. J. Conejo, J. Morales, and L. Baringo, "Real-time demand response model," *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 236 – 242, Dec. 2010.

[61] P. Faria, Z. Vale, J. Soares, and J. Ferreira, "Demand response management in power systems using particle swarm optimization," *IEEE Intelligent Systems*, vol. 28, no. 4, pp. 43–51, July 2013.

[62] K. M. Tsui and S. C. Chan, "Demand response optimization for smart home scheduling under real-time pricing," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1812–1821, Dec. 2012.

[63] S. H. Madaeni and R. Sioshansi, "Measuring the benefits of delayed price-responsive demand in reducing wind-uncertainty costs," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4118–4126, Nov. 2013.

[64] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2008.

[65] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & sons, 2008.

[66] K. Mets, T. Verschueren, W. Haerick, C. Develder, and F. D. Turck, "Optimizing smart energy control strategies for plug-in hybrid electric vehicle charging," in *IEEE/IFIP NOMS Wksps*, Osaka, Japan, Apr. 2010, pp. 293 – 299.

[67] E. Muljadi and A. Ellis, "Final project report wecc wind generator development appendix v: Model validation of wind turbine generator," National Renewable Energy Laboratory, Tech. Rep., Mar. 2010.

[68] L. H. Tsoukalas and R. Gao, "From smart grids to an energy internet: Assumptions, architectures and requirements," in *Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, Nanjing, China, Apr. 2008, pp. 94 –98.

[69] R. N. Anderson, A. Boulanger, W. B. Powell, and W. Scott, "Adaptive stochastic control for the smart grid," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 1098 –1115, June 2011.

[70] A. Battaglini, J. Lilliestam, A. Haas, and A. Patt, "Development of supersmart grids for a more efficient utilisation of electricity from renewable sources," *Journal of Cleaner Production*, vol. 17, no. 10, pp. 911 – 918, 2009.

[71] P. P. Varaiya, F. F. Wu, and J. W. Bialek, "Smart operation of smart grid: Risk-limiting dispatch," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 40 –57, Jan. 2011.

[72] S. K. Jayaweera and D. Li, "Stochastic control for smart grid with integrated renewable distributed generators," University of New Mexico, Tech. Rep., Jan. 2011.

[73] D. Li, S. K. Jayaweera, and C. T. Abdallah, "Uncertainty modeling and stochastic control design for smart grid with distributed renewables," in *2012 IEEE Green Technologies Conference*, Tulsa, OK, USA, Apr. 2012.

[74] D. Li and S. K. Jayaweera, "Stochastic control for utility-maintained smartgrid with integrated synchronous rdg's," in 2013 *IEEE Online Conference on Green Communications (GreenCom)*, Nov. 2013, in review.

[75] Y.-N. Yu, *Eletric Power System Dynamics*. New York: Academic, 1983.

[76] M. Dehghani and S. Nikravesh, "Nonlinear state space model identification of synchronous generators," *Electric Power Systems Research*, vol. 78, no. 5, pp. 926 – 940, 2008.

[77] F. Kupzog, T. Sauter, and KlausPollhammer, "It-enabled integration of renewables: a concept for the smart power grid," *EURASIP J. Embedded Syst.*, vol. 2011, pp. 5:1–5:8, January 2011.

[78] J. A. P. Lopes, P. M. R. Almeida, and F. J. Soares, "Using vehicle-to-grid to maximize the integration of intermittent renewable energy resources in islanded electric grids," in *2009 International Conference on Clean Electrical Power*, Jun. 2009, pp. 290 –295.

[79] G. D. Rodriguez, "A utility perspective of the role of energy storage in the smart grid," in *2010 IEEE Power and Energy Society General Meeting*, Jul. 2010, pp. 1 –2.

[80] M. Wang and Y. Sun, "A practical method to improve phasor and power measurement accuracy of dft algorithm," *IEEE Transactions on Power Delivery*, vol. 21, no. 3, pp. 1054 – 1062, Jul. 2006.

[81] L. Durnte and P. K. Ghosh, "Active power measurement in nonsinusoidal environments," *Power Systems, IEEE Transactions on*, vol. 15, no. 3, pp. 1142 – 1147, Aug. 2000.

[82] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. New York: Wiley Interscience, 1972.

[83] N. Sharma, P. Sharma, D. Irwin, and P. Shenoy, "Predicting solar generation from weather forecasts using machine learning," in *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Oct. 2011, pp. 528 –533.

[84] D. S. Naidu, *Optimal Control Systems*. Boca Raton, Florida, USA: CRC Press, 2003.

[85] J. Lopes, N. Hatziargyriou, J. Mutale, P. Djapic, and N. Jenkins, "Integrating distributed generation into electric power systems: A review of drivers, challenges and opportunities," *Electric Power Systems Research*, vol. 77, no. 9, pp. 1189–1203, Jul. 2007.

[86] H. V. Poor, *An introduction to signal detection and estimation (2nd ed.)*. New York, NY, USA: Springer-Verlag New York, Inc., 1994.

[87] H. S. Cho, T. Yamazaki, and M. Hahn, "Determining location of appliances from multi-hop tree structures of power strip type smart meters," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 4, pp. 2314 – 2322, Nov. 2009.

[88] W. Saad, Z. Han, H. V. Poor, and T. Basar, "Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 86–105, 2012.

[89] L. Huang, J. Walrand, and K. Ramchandran, "Optimal demand response with energy storage management," in *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, 2012, pp. 61–66.

[90] ——, "Optimal power procurement and demand response with quality-of-usage guarantees," in *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1–8.

[91] A. Lam, L. Huang, A. Silva, and W. Saad, "A multi-layer market for vehicle-to-grid energy trading in the smart grid," in *2012 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS),*, 2012, pp. 85–90.

[92] M. Parvania and M. Fotuhi-Firuzabad, "Demand response scheduling by stochastic scuc," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 89 – 98, Jun. 2010.

[93] A. H. Mohsenian-Rad and A. Leon-Garcia, "Optimal residential load control with price prediction in real-time electricity pricing environments," *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 120 – 133, Sept. 2010.

[94] D. Li and S. K. Jayaweera, "Optimal stochastic tracking for primary frequency control in an interactive smart grid infrastructure," *IEEE Systems Journal*, Apr. 2014, in review.

[95] S. El-Ferik and R. P. Malhame, "Identification of alternating renewal electric load models from energy measurements," *IEEE Transactions on Automatic Control*, vol. 39, no. 6, pp. 1184 –1196, Jun. 1994.

[96] E. Bellone, J. P. Hughes, and P. Guttorp, "A hidden markov model for downscaling synoptic atmospheric patterns to precipitation amounts," *Climate research*, vol. 15, no. 1, pp. 1–15, May 2000.

[97] D. Li and S. K. Jayaweera, "Uncertainty modeling and prediction for customer load demand in smart grid," in 2013 *IEEE EnergyTech Conference*, Cleveland, Ohio, USA, May 2013.

[98] A. R. Cassandro, "Exact and Approximate Algorithms for Partially Observable Markov Decision Processes," Ph.D. dissertation, Brown University, 1998.

[99] W. B. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2011.

[100] A. W. Moore and C. G. Atkeson, "Prioritized sweeping: Reinforcement learning with less data and less real time," *Machine Learning*, vol. 13, pp. 103 – 130, Oct. 1993.

[101] L. F. Bertuccelli and J. P. How, "Estimation of non-stationary markov chain transition models," in *the 47th IEEE Conference on Decision and Control (CDC 2008)*, Cancun, Mexico, Dec. 2008, pp. 55 –60.

[102] R. Jaulmes, J. Pineau, and D. Precup, "Learning in non-stationary partially observable markov decision processes," in *ECML Workshop on Reinforcement Learning in Non-Stationary Environments*, Porto, Portugal, Dec. 2005, pp. 55 – 60.

[103] J. Y. Yu and S. Mannor, "Arbitrarily modulated markov decision processes," in *Proceedings of the 48th IEEE Conference on Decision and Control*, Dec. 2009, pp. 2946 –2953.

[104] S. P. M. Choi, D.-Y. Yeung, and N. L. Zhang, "Hidden-mode markov decision processes for nonstationary sequential decision making," in *In Sequence Learning - Paradigms, Algorithms, and Applications*. Springer-Verlag, 2001, pp. 264–287.

[105] L. R. Welch, "Hidden Markov Models and the Baum-Welch Algorithm," *IEEE Information Theory Society Newsletter*, vol. 53, no. 4, Dec. 2003.

[106] S. C. Dit-Yan, S. P. M. Choi, D. yan Yeung, and N. L. Zhang, "Hidden-mode markov decision processes," in *In IJCAI Workshop on Neural, Symbolic, and Reinforcement Methods for Sequence Learning*, 1999.

[107] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, Massachusetts: The MIT Press, 1998.

[108] M. L. Littman, "Algorithms for Sequential Decision Making," Ph.D. dissertation, Stanford University, 1996.

[109] A. Cassandra, M. L. Littman, and N. L. Zhang, "Incremental pruning: A simple, fast, exact method for partially observable markov decision processes," in *In Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann Publishers, 1997, pp. 54–61.

[110] G. E. Monahan, "A survey of partially observable markov decision processes: Theory, models, and algorit," *Management Science*, vol. 28, no. 1, pp. 1–16, Jan. 1982.

[111] E. J. Sondik, "The optimal control of partially observable Markov processes over the infinite horizon: Discounted costs," *Operations Research*, vol. 26, pp. 282–304, 1978.

[112] M. L. Littman, "The witness algorithm: Solving partially observable markov decision processes," Brown University, Tech. Rep., 1994.

[113] R. Soultanian, "Global electricity pricing: Ups and downs of global electricity prices," *Power Engineering International Magazine*, Jul. 2007.

[114] C. J. C. H. Watkins, "Learning from Delayed Rewards," Ph.D. dissertation, Cambridge University, Cambridge, England., 1989.

[115] F. S. Melo and M. I. Ribeiro, "Q-learning with linear function approximation," in *Proceedings of the 20th Annual Conference on Learning Theory*. Springer-Verlag, 2007, pp. 308–322.

[116] V. Krishna, *Auction Theory*. Academic Press, Elsevier, 2010.

[117] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, 2010.

[118] D. Fudenberg and J. Tirole, *Game Theory*. MIT Press, 1991.

[119] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*. Cambridge University Press, 2007.

[120] A. Blume, P. Heidhues, J. Lafky, J. Munster, and M. Zhang, "All equilibria of the multi-unit vickrey auction," *Games and Economic Behavior*, vol. 66, no. 2, pp. 729–741, 2009.

[121] A. Blume and P. Heidhues, "All equilibria of the vickrey auction," *Journal of Economic Theory*, vol. 114, pp. 170–177, 2004.

[122] D. Bernstein, R. Givan, N. Immerman, and S. Zilberstein, "The complexity of decentralized control of markov decision processes," *Mathematics of Operations Research*, vol. 27, no. 4, pp. 819–840, 2002.