REAL-TIME PRICING IN SMART GRID: IMPACTS OF DATA, INTERACTIONS AND LEARNING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Liyan Jia

May 2017

REAL-TIME PRICING IN SMART GRID: IMPACTS OF DATA, INTERACTIONS AND LEARNING

Liyan Jia, Ph.D.

Cornell University 2017

In modern deregulated electricity market, price is playing a key role in signaling the economic benefits to all participants. In particular, real-time price dynamically reflects the actual marginal cost of generating electricity and directly associates with market settlement. Methods of quantifying the effect of real-time pricing and optimal pricing policy design are needed to ensure reliable system operation and improve economic efficiency. In this thesis, three problems associated with real-time pricing in smart grid are studied.

First, the impacts of data quality on real-time locational marginal price (LMP) is characterized. Because the real-time LMP is computed from the estimated network topology and system state, bad data that cause errors in topology processing and state estimation affect real-time LMP. It is shown that the power system state space is partitioned into price regions of convex polytopes. Under different bad data models, the worst case impacts of bad data on real-time LMP are analyzed. Numerical simulations are used to illustrate worst case performance for IEEE-14 and IEEE-118 networks.

Second, the problem of designing dynamic price of electricity in retail market is considered. For the day-ahead hourly pricing (DAHP) scheme, a Stackelberg game model is formulated with the retailer as a leader and its customers as followers. By solving a real-time load control problem, an affine structure between the optimal demand response and day-ahead retail price is established and the

trade-off curve between consumer surplus (CS) and retail profit (RP) is characterized as a concave Pareto front, on which each point is an equilibrium of the Stackelberg game with a particular retailer's payoff function. Effects of renewable energy and storage are also analyzed under the same Stackelberg game model. It is shown that the tradeoff curves with renewable energy or storage on the retailer side have significantly different characteristics from those ones on the consumer side. Simulations based on actual weather and price data are used to verify our statements.

Finally, the problem of optimal dynamic pricing for retail electricity with an unknown demand model is considered. Without knowledge on the aggregated demand function of its customers, a retailer aims to maximize its retail surplus by sequentially adjusting its price based on the behavior of its customers in the past. An online learning algorithm, referred to as piecewise linear stochastic approximation (PWLSA), is proposed. It is shown that PWLSA achieves the optimal rate of learning defined by the growth rate of cumulative regret. In particular, the regret of PWLSA is shown to grow logarithmically with respect to the learning horizon, and no other on-line learning algorithm can have the growth rate slower than that of PWLSA. Simulation studies are presented using traces of actual day-ahead prices, and PWLSA compares favorably under both static and dynamically changing parameters.

BIOGRAPHICAL SKETCH

Liyan Jia was born in Handan, Hebei, China in 1987. He graduated in 2005 with bachelor's degree in Department of Automation at Tsinghua University, Beijing. In August 2009, he became a PhD student at Cornell University, and joined the Adaptive Communications and Signal Processing (ACSP) group under Prof. Lang Tong. His research mostly focuses on electricity market and dynamic programming. In his spare time, Liyan is a big fan of soccer and also enjoys watching movies, playing chess, and solving puzzles.

For my parents, Guochen Jia and Shu'ai Cui

ACKNOWLEDGEMENTS

Above all, I would like to thank my advisor, Lang Tong, who is like the lighthouse guiding me through the past five years. I'll never forget those sparks initiated during our discussion, those encouraging words he said when I was depressed, and those busy days and nights we worked together on slides and papers. He opens a new research world to me and always stands behind me being supportive. His enthusiasm, his way of critical thinking and his sense of humor will influence my life forever.

I also want to thank the other members of my thesis committee. Hsiao-dong Chiang, for teaching me two basic power system courses and being a constant guide on my research adventure. Timothy Mount, for knowing everything about economics and power market.

I want to thank the former and current members in our research group: Oliver Kosut, Brandon M. Jones, Shiyao Chen, Jinsub Kim, Zhe Yu, Yuting Ji, Daniel Munoz Alvarez, and Tirza Routtenberg. They have provided me with helpful research discussions, personal supports, and lots of small talks. I thank them for all the fun we had together.

I also want to thank other professors, who taught me at Cornell, for giving wonderful lectures: Prof. Aaron B. Wagner, Prof. Gennady Samorodnitsky, Prof. Sidney Ira Resnick, and Prof. Michael Jeremy Todd.

I would like to thank my parents, Guochen Jia and Shu'ai Cui. They have helped me all the way since February 1987. I thank them for everything I am now.

This thesis is supported in part by the National Science Foundation under Grant CNS-1135844.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Overview and motivation

In modern deregulated electricity market, price is playing a key role in signaling the economic benefits to all participants. In particular, real-time price dynamically reflects the actual marginal cost of generating electricity and directly associates with market settlement. Methods of quantifying the effect of real-time pricing and optimal pricing policy design are needed to ensure reliable system operation and improve economic efficiency. In this thesis, we studied both the wholesale and retail electricity market, focusing on data quality for the former and dynamic retail pricing design for the latter.

1.1.1 Wholesale market

Most deregulated wholesale electricity markets in U.S. have two interconnected components. The day-ahead market determines the locational marginal price (LMP) based on the dual variables of the optimal power flow (OPF) solution [68, 45], given generator offers, demand forecast, system topology, and security constraints. The calculation of LMP in the day-ahead market does not depend on the actual system operation. In the real-time market, on the other hand, an ex-post formulation is often used (*e.g.*, by PJM and ISO-New England [73]) to calculate the real-time LMP by solving an incremental OPF problem. The LMPs in the day-ahead and the real-time markets are combined in the final clearing and settlement processes.

The real-time LMP is a function of data collected by the supervisory control and data acquisition (SCADA) system. Therefore, anomalies in data, if undetected, will affect prices in the real-time market. While the control center employs a bad data detector to "clean" the real-time measurements, miss detections and false alarms will occur inevitably. The increasing reliance on the cyber system also comes with the risk that malicious data may be injected by an adversary to affect system and real-time market operations. An intelligent adversary can carefully design a data attack to avoid detection by the bad data detector.

Regardless of the source of data errors, it is of significant value to assess potential impacts of data quality on the real-time market, especially when a smart grid may in the future deploy demand response based on real-time LMP. To this end, we are interested in characterizing the impact of worst case data errors on the real-time LMP. The focus on the worst case also reflects the lack of an accurate model of bad data and our desire to include the possibility of data attacks.

1.1.2 Demand response and retail market

As a key feature of a future smart grid, demand response is expected to offer economic benefits to the consumers by taking advantage of features of a smart grid and intelligent pricing mechanisms. Through various mechanisms, demand response can reshape the customers' consumption patterns. Properly designed demand response program offers the potential to reduce the peak hour energy usage, which gives system operator flexibility to improve the operating

efficiency and reliability, compensate the uncertainty introduced by renewable integration, and in the end increase the social welfare.

According to a report of FERC (Federal Energy Regulatory Commission) [20], existing demand response programs can be grouped into two categories, incentive-based demand response and time-based rates. The incentive-based programs offer participating consumers reduced price or some guaranteed economic benefits through a long term contract, in exchange for some level of control power over the consumers' energy usage. Examples[20] of such programs include direct load control (where the consumers' electric equipment can be shut down or cycled remotely), interruptible/curtailable rates (where electric consumption will be curtailed in case of emergency), demand bidding/buyback programs (where dispatchable demand resource can participate in market bidding), and etc. Although these programs give the operator the flexibility to optimize system operating decisions or handle contingencies, a consumer loses the ability to manage her own energy usage based on her own preference.

Alternatively, time-based rates use pricing mechanisms to affect the consumers' energy usage indirectly. In these programs, a consumer is empowered to fully optimize her consumption, taking into account her needs and the price signals provided by the retailer. This kind of programs makes more economic sense and avoids many legal problems. Examples[20] include time-of-use rates (where prices are pre-set for predetermined periods in advance), critical-peak pricing(where prespecified high rates are imposed for a limited number of days or hours in case of wholesale price peak or contingency), and real-time pricing (retail price fluctuates hourly or more often, to reflect changes in the wholesale price on either a day-ahead or hour-ahead basis).

The advent of a smart grid makes the time-based rates demand response potentially more attractive. By taking advantage of the enhanced sensing, communication, and computation power at the consumer side, distributed energy management allows the consumers to optimize individualized trade-offs between energy cost and the quality of service.

Currently, the structure of most electricity retail markets is monopoly and the retail price is highly regulated. However, demand response programs and wholesale market deregulation present nontrivial challenges to cope with wholesale price fluctuations and dynamically managed consumer demand at the same time. It is interesting to see how the retailer interact with its consumers, via the dynamically designed retail price.

1.1.3 Renewable integration and demand response

Renewable integration in distribution networks is undergoing a phenomenal growth. According to the latest data [1], the cumulative residential photovoltaic (PV) installation grows from 1.1 GW in the first half of 2012 to 2.7GW in the first half of 2014 with a 45% increase in the second quarter of 2014 over the same quarter in 2013. The cumulative installation of utility PV is quadrupled from the 2012, reaching the level of 7 GW in 2014. These statistics suggests an accelerated participation in renewable integration by both utilities and consumers.

The economics of renewable integration by the utility and that by individual consumers are fundamentally different, even though both reduce the amount of energy drawn from the transmission grid. For a consumer based renewable integration, there is a current debate on the use of net metering that allows a

consumer sell surplus electricity generated from its solar PV back to the utility at the same price. It is apparent that net metering undermines the revenue of the utility, making it difficult for the utility to maintain and invest in the distribution infrastructure.

Therefore, the effect of renewable integration on the interaction between utility and consumers is nontrivial and whether the consumers will benefit from renewable integration or not becomes unclear.

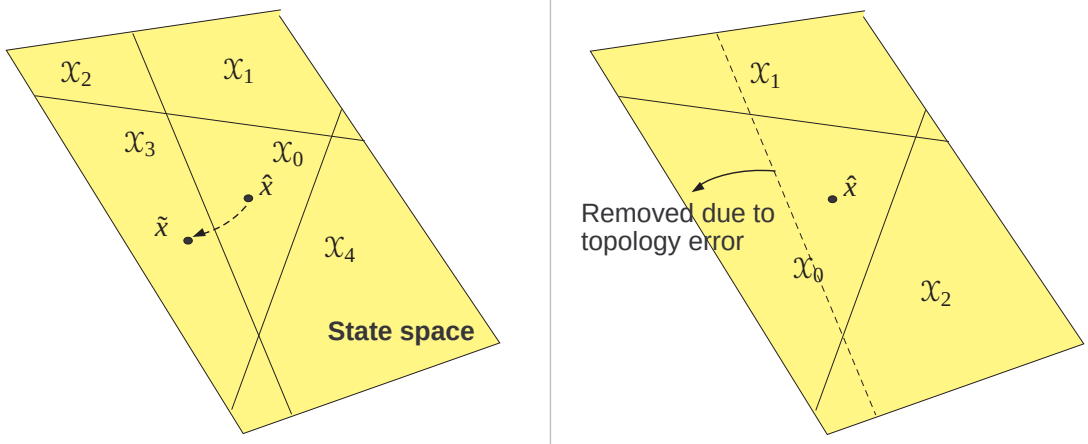
1.2 Effect of data quality on real-time wholesale price

In this part, we aim to characterize the worst effects of data corruption on real-time LMP. By "worst", we mean the maximum perturbation of real-time LMP caused by bad or malicious data, when a fixed set of data is subject to corruption. The complete characterization of worst data impact, however, is not computationally tractable. Our goal here is to develop an optimization based approach to search for locally worst data by restricting the network congestion to a set of lines prone to congestion. We then apply computationally tractable (greedy search) algorithms to find the worst data and evaluate the effects of worst data by simulations.

1.2.1 Summary of results

In characterizing the relation between data and real-time LMP, we first present a geometric characterization of the real-time LMP. In particular, we show that the state space of the power system is partitioned into polytope price regions,

as illustrated in Fig. 1.1(a), where each polytope is associated with a unique real-time LMP vector, and the price region $\mathcal{X}_i$ is defined by a particular set of congested lines that determine the boundaries of the price region.



(a) Bad meter data (b) Topology error

Figure 1.1: Change of real-time LMPs due to bad data.

Two types of bad data are considered. One is the bad data associated with meter measurements such as the branch power flows in the network. Such bad data will cause errors in state estimation, possibly perturbing, as an example, the correct state estimate $\hat{x}$ in $\mathcal{X}_0$ to $\tilde{x}$ in $\mathcal{X}_3$ (as shown in Fig. 1.1(a)). The analysis of the worst case data then corresponds to finding the worst measurement error such that it perturbs the correct state estimation to the worst price region.

The second type of bad data, one that has not been carefully studied in the context of LMP in the literature, is error in digital measurements such as switch or breaker states. Such errors lead directly to topology errors therefore causing a change in the polytope structure as illustrated in Fig. 1.1(b). In this case, even if the estimated system state changes little, the prices associated with each region change, sometimes quite significantly.

Before characterizing impacts of bad meter data on LMP, we need to construct appropriate models for bad data. To this end, we propose three increasingly more powerful bad data models based on the dependencies on real-time system measurements: state independent bad data, partially adaptive bad data, and fully adaptive bad data.

In studying the worst case performance, we adopt a widely used approach that casts the problem as one involving an adversary whose goal is to make the system performance as poor as possible. The approach of finding the worst data is equivalent to finding the optimal strategy of an attacker who tries to perturb the real-time LMP and avoid being detected at the same time. By giving the adversary more information about the network state and endowing him with the ability to change data, we are able to capture the worst case performance, sometimes exactly and sometimes as bounds on performance.

Finally, we perform simulation studies using the IEEE-14 and IEEE-118 networks. We observe that bad data independent of the system state seems to have limited impact on real-time LMPs, and greater price perturbations can be achieved by state dependent bad data. The results also demonstrate that the real-time LMPs are subject to much larger perturbation if bad topology data are present in addition to bad meter data.

While substantial price changes can be realized for small networks by the worst meter data, as the size of network grows while the measurement redundancy rate remains the same, the influence of worst meter data on LMP is reduced. However, larger system actually gives more possibilities for the bad topology data to perturb the real-time LMP more significantly.

Our simulation results also show a degree of robustness provided by the nonlinear state estimator. While there have been many studies on data injection attacks based on DC models, very few consider the fact that the control center typically employs the nonlinear state estimator under the AC model. Our simulation shows that effects of bad analog data designed based on DC model may be mitigated by the nonlinear estimator whereas bad topology data coupled with bad analog data can have greater impacts on LMP.

1.2.2 Related Work

Effects of bad data on power system have been studied extensively in the past, see [5, 23, 60]. Finding the worst case bad data is naturally connected with the problem of malicious data. In this context, the results presented here can be viewed as one of analyzing the impact of the worst (malicious) data attack.

In a seminal paper by Liu, Ning, and Reiter [46], the authors first illustrated the possibility that, by compromising enough number of meters, an adversary can perturb the state estimate arbitrarily in some subspace of the state space without being detected by any bad data detector. Such attacks are referred to as strong attacks. It was shown by Kosut *et al.* [38] that the condition for the existence of such undetectable attacks is equivalent to the classical notion of network observability.

When the adversary can only inject malicious data from a small number of meters, strong attacks do not exist, and any injected malicious data can be detected with some probability. Such attacks are referred to as weak attacks [38]. In order to affect the system operation in some meaningful way, the adversary

has to risk being detected by the control center. The impacts of weak attack on power system are not well understood because the detection of such bad data is probabilistic. Our results are perhaps the first to quantify such impacts. Most related research works focused on DC model and linear estimator while only few have addressed the nonlinearity effect [28, 27].

It is well recognized that bad data can also cause topology errors [67, 16], and techniques have been developed to detect topology errors. For instance, the residue vector from state estimation was analyzed for topology error detection [16, 67, 19]. Monticelli [50] introduced the idea of generalized state estimation where, roughly speaking, the topology that fits the meter measurements best is chosen as the topology estimate. The impacts of topology errors on electricity market have not been reported in the literature, and our work aims to bridge this gap.

The effect of data quality on real-time market was first considered in [66, 69]. In [69], the authors presented the financial risks induced by the data perturbation and proposed a heuristic technique for finding a case where price change happens. While there are similarities between our work and [69], several significant differences exist: (i) We focus on finding the worst case, not only a feasible case. (ii) Our work considers a more general class of bad data where bad data may depend dynamically on the actual system measurements rather than static. (iii) We consider a broader range of bad data that also include bad topology data, and our evaluations are based on the AC network model and the presence of nonlinear state estimator.

1.3 Interactions between electricity retailer and consumers

In this part, we consider a demand response scheme based on price incentives provided by the retailer and distributed energy management by its customers. In particular, we focus on a specific pricing scheme, named day-ahead hourly pricing (DAHP), where the hourly prices are posted one day ahead by the retailer, while the payment is settled in real time as the product of the consumers' dynamic electricity usage and the fixed day-ahead hourly prices.

The advantage of DAHP is twofold. For a consumer, DAHP gives her the price certainty one day ahead of time so that she can plan accordingly based on the posted prices and her desired quality of service. On the other hand, DAHP allows the retailer to adjust its prices on a day-to-day basis, taking into account operating conditions at the wholesale market (such as the day-ahead wholesale price) and environmental factors (such as temperatures and projected renewable generations). A central theme here is therefore the complete characterization of the DAHP induced tradeoffs between consumer surplus and retail profit.

Integrating renewable energy into distributed networks makes the problem more complicated. Is consumer based renewable integration good for consumers? The answer may be less obvious than it appears at the first glance. When the utility is a regulated monopoly, the loss of demand from the consumer may force the utility to raise the price of electricity, which affects the consumer surplus. If the consumer's capacity of renewable integration is not sufficiently high to have net zero consumption, he may be worse off than he is in absence of local renewable integration. Such schemes also place unproportionally heavy burden on those without privately installed renewable generation.

What happens for the centralized renewable integration owned by or contracted to a utility? The access to renewable by the utility reduces the amount of power procured from the wholesale market and allows the utility hedge against real-time price spikes in the wholesale market. Will the utility transfer at least part of the surplus from the renewable integration back to its consumers? For example, will both the utility and the end consumer benefit if the utility provides a rate discount to the consumer by operating and maintaining consumer's roof top PV installations?

1.3.1 Summary of results

We present a Stackelberg game model to analyze and optimize a DAHP based demand response, with the retailer as a leader and the consumers as followers. For HVAC* demands, the optimal demand response is formulated by solving a distributed stochastic program that maximizes the consumer surplus. The optimal solution to the program reveals an affine structure of the demand function. For the retailer, on the other hand, under various payoff functions, the optimal DAHPs are obtained via convex optimizations.

We provide a complete characterization of the Pareto front of the trade-off between consumer surplus (CS) and retail profit (RP), as illustrated in Fig. 1.2. We show that the Pareto front that characterizes the optimal tradeoff is given by a concave and monotonically decreasing function of the consumer surplus. It consists of Stackelberg equilibria achievable by optimized DAHP; any CS-RP pair above the Pareto front is not attainable.

*heating, ventilation, and air conditioning

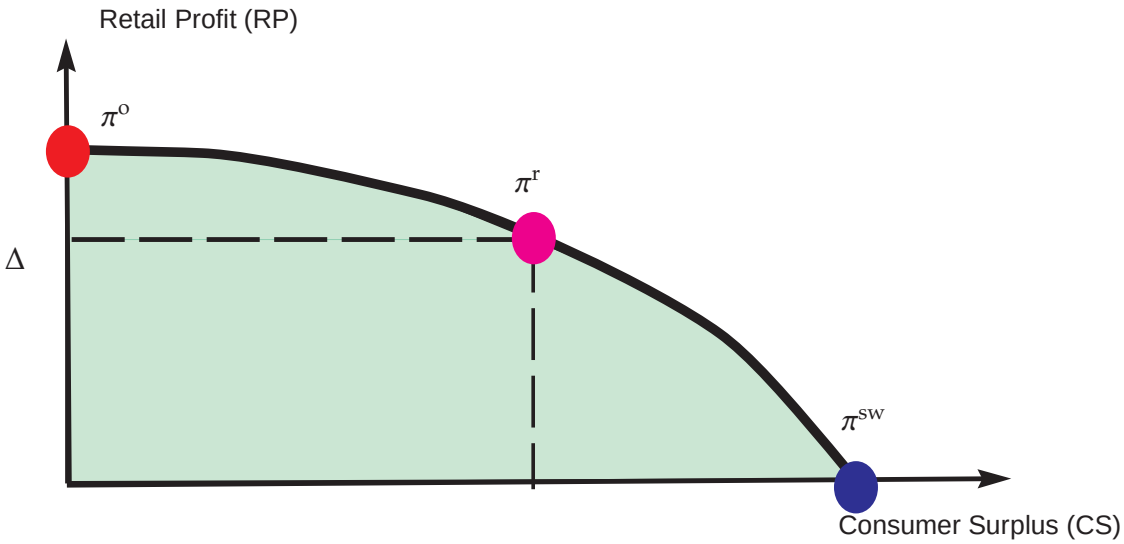


Figure 1.2: CS-RP trade-off curve with different dynamic pricing schemes

We can place several well known pricing schemes on the CS-RP plane. In particular, the social welfare maximizing price π^{sw} is shown to result in zero retail profit, which means that social welfare maximization is not economically viable to the retailer. The optimal regulated monopoly price π^r is located at the Pareto front where the retailer profit has a regulated profit margin Δ. The retailer profit maximization price is π^o when no constraint is imposed on the retailer. We note also that benchmark pricing schemes such as constant pricing and proportional markup pricing are strictly inside Pareto front, indicating that benefits can be improved for both the consumers and the retailer.

We also consider the effects of incorporating renewable energy and storage devices on either the utility side or the consumer side. We are interested in particular the change of the Pareto front when as more renewable sources are

utilized or higher capacity of storage is made available.

We compare two models of renewable generation: the first is a *centralized utility-based model* in which the utility owns and operates the renewable generation as part of its portfolio of energy resources. The second is a *decentralized consumer-based model* in which a consumer owns and operates the renewable generation and is allowed to sell back surplus electricity in a net-metering setting. These models are of course stylized abstractions that differ from practical implementations; the two models coexist in practice as indicated by the recent PV installation statistics [1]. However, the main features that underline the debate on net metering are retained.

To analyze effects of renewable integration on the utility and the consumer, we characterize the Pareto frontier in the tradeoff between the retail profit of the utility and the consumer surplus. Because the Pareto frontier specifies the maximum achievable consumer surplus for a fixed retail profit, we are able to compare the corresponding consumer surpluses from different integration models by analyzing their associated Pareto frontiers.

We obtain the following analytical results using a Stackleberg game analysis under the net metering mechanism where the prices of purchasing from and selling to the utility are the same. A graphical sketch of the results is shown in Fig. 1.3.

Centralized utility-based integration For the utility-based renewable integration, as expected, the Pareto frontier is strictly above that when there is no re-

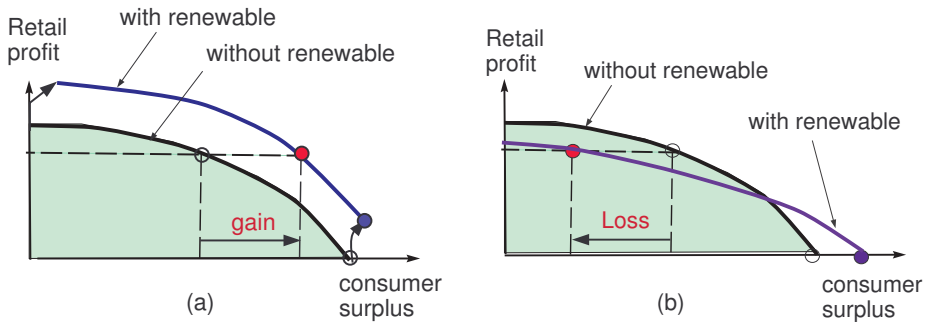


Figure 1.3: Trade-off curves with renewable integration: (a) Utility based renewable integration. (b) Consumer-based renewable integration.

newable integration, as shown in Fig. 1.3(a). Our result, however, quantifies the benefit and shows how the benefit of renewable integration is distributed between the utility and the consumers. In particular, we show in Theorem 6 that, the retailer distributes the benefit of wind integration only if the wind power capacity K exceeds a certain threshold. As K increases, the benefits from wind integration apportioned to consumer increases monotonically. The implication is that, for a regulated monopoly, *the consumer always benefits indirectly from centralized renewable integration by the utility in the sense that its consumer surplus always is increased.*

Decentralized consumer-based integration For the consumer-based integration, we show that in Theorem 7 the Pareto frontier intersects with the Pareto frontier when there is no renewable integration, as illustrated in Fig. 1.3(b). This means that, *if the regulated profit is to be maintained by the monopolistic utility, the consumer may wind up with a lower level consumer surplus.* In other words, consumer owned and operated renewable may be bad for consumers. This perhaps is unexpected at first the glance; it does confirm the intuition that the reduction of consumption forces the regulated utility to increase the price of electricity,

which lowered the consumer surplus.

Similar results are also shown for storage devices. For the case that energy storage device is used by the utility, we show that the trade-off curve is simply shifted up by the amount of arbitrage profit over the wholesale price, as shown in Fig 1.4. On the other hand, when the storage device is owned by the consumers, the consumers' HVAC control policy remains the same while the total demand is changed purely by the amount of storage control. The consumer surplus is increased by the arbitrage profit over retail price. Theorem 8 proves that with consumer-based storage, the maximized social welfare will increase while the maximized profit will decrease under some conditions. Therefore, the new trade-off curve will cross with the original one, as shown in Fig 1.4. Theorem 8 also shows that if the utility-based storage has the same parameters as the aggregated consumer-based storage, the trade-off curve with utility-based storage will be completely outside the one with consumer-based storage. This shows that centralized control is more efficient than the distributed control under this monopoly retail market structure. Simulations of the effect of energy storage devices are presented in Fig. 3.6 and Fig. 3.7 in Section 3.5.5.

Numerical results for consumers with thermal load are using more realistic parameters to confirm the analytical results.

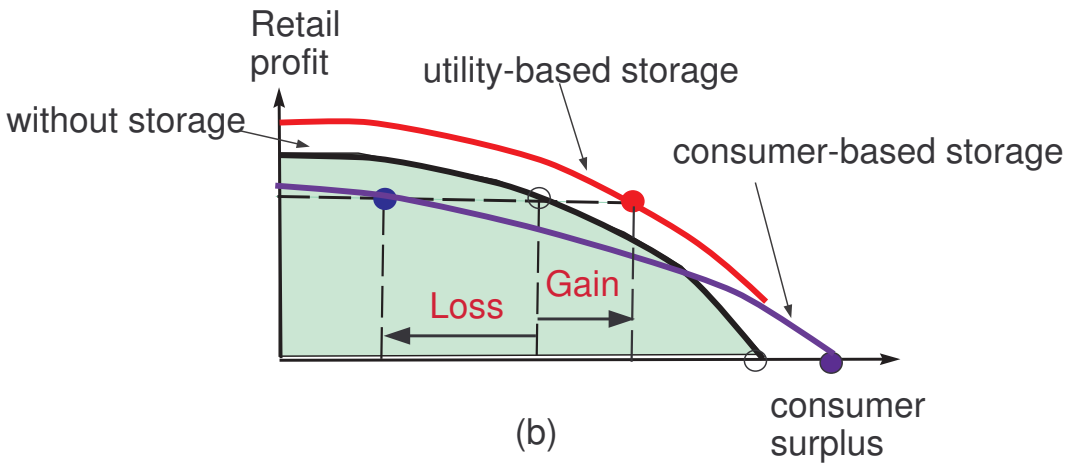


Figure 1.4: Trade-off curves with storage

1.3.2 Related work

The study of dynamic electricity pricing at retail level appeared along with the beginning of wholesale market deregulation. In [12, 11], based on a simplified model and simulations, the authors demonstrate the benefit of introducing dynamic pricing, including economic incentive to the demand side and long-run operation efficiency. Practical programs of dynamic pricing, including DAHP considered here, have been implemented by the utility companies among U.S., such as Niagara Mohawk [25], Georgia Power, Gulf Power and etc. By empirical study, the authors of [25], where the name of DAHP is adopted from, conclude that DAHP "not only improves the linkage between wholesale and retail markets, but also promotes the development of retail competition." The effect and benefit of DAHP have also been studied theoretically in [12], which shows that DAHP provides more benefits to consumers and can attract, in the long run, all the consumers, compared with other pricing schemes such as flat rate and time of use (TOU).

However, those schemes are either ad-hoc or just passing the wholesale price with some mark-up to the consumers, without any optimization as considered in our work. In [15, 17], a comprehensive framework of retail price optimization is formulated, by considering the consumer side uncertainty and the resulted financial risk. Relative to this line of existing work, the main contribution of our work is the optimized DAHP, the use of a Stackelberg game model to obtain a full characterization of the trade-off between the achievable retail profit and consumer surplus, and the effect of renewable energy on the trade-off curve.

In our work, a Stackelberg game model is used to formulate the relationship between a retailer and its consumers. A similar idea can be found in [48], which finds the optimal retail price to induce social welfare maximization consumption under a repeated Stackelberg game model. [71] also uses a game theoretical setting to study the interactive retail market structure, based on an abstract and static model for the price responsive demand.

Extensive research work has been done focusing on demand response optimization. Several recent studies are using a distributed fashion to find the social welfare maximization consumption, where retail price is an ancillary variable in each iteration and a by-product after the optimization. For example, [31] designs a hierarchical market structure that includes the loads at the bottom layer, load service entities at the middle layer, and a independent service operator at the top. By exchanging decisions between different layers in each iteration, a global solution will be achieved with guaranteed convergence, while the operator dose not need to know full information about the appliances in the bottom layer. Focusing on the retail level optimization, a similar distributed algorithm is used to achieve social welfare maximization in [43]. The contract design be-

tween the renewable generators and the aggregator is studied in [54], where the aggregator is responsible for a large scale PHEV charging.

Renewable integration and usage of energy storage are also among the recent hot topics in smart grid research, especially with the help of demand response programs, while most of them are without addressing the benefit-cost tradeoff in the interactions between the utility and consumers. A relevant work is the contract design between the renewable generators and the aggregator studied in [54], where the aggregator is responsible for a large scale PHEV charging. The work in [53] shows that by coupling with deferrable loads, the costs resulting from the randomness of wind power supply can be significantly mitigated, which makes large scale wind power integration possible without the need of significant investments in backup generation. Relevant to our work, [14] shows that electricity rates must increase for utility companies to recover their fixed costs when PVs are implemented while higher electricity rates give households more incentives to adopt PVs based on constant demand profiles. For consumer-based energy storage, [70] formulates an stochastic optimization problem and characterizes the properties of economic value of energy storage. Similar ideas can also be found in [62, 39]. Distinct from the topics in those papers, to our knowledge, our work is the first to study the effect of renewable energy and storage on the interaction between retailer and consumers. See also *e.g.,* [32, 61] on the constraints and allocation of embedded renewable sources, [63] on the impact of renewable on the distribution systems, [26] on pricing and congestion management, and [59, 64] on the control and optimization of distributed generation. There is a substantial literature on optimal demand response. See recent work in [44, 31] and references therein.

1.4 Online learning of dynamic retail price

As we discussed above in Section 1.3, the consumer's demand is a affine function of the retail price. If a retailer knows how its customers respond to the retail price through their individual demand functions, it can choose the price to optimize a particular objective, *e.g.*, the social welfare or its own profit subject to regulations. Obtaining the demand functions of its customers, however, is nontrivial because a customer is likely to consider such information private; neither the willingness of sharing nor the correctness of the shared information can be assumed.

Our work focuses on optimal dynamic pricing under unknown demand functions. We take an online learning approach where the retailer learns the behavior of its customers by observing their response to carefully designed prices. The basic principle of online learning is to achieve a tradeoff between "exploration" and "exploitation;" the former represents the need of using sufficiently rich pricing signals to achieve learning accuracy, whereas the latter stands for the need of capturing as much reward as possible based on what has been learned.

In the classical online learning theory, the performance of a learning algorithm is measured by the notion of cumulative regret. For the pricing problem at hand, the regret is defined as the difference between the retail surplus associated with the actual aggregated demand function and the surplus achieved by an online learning algorithm. While the cumulative regret R_T grows with the learning horizon T, the rate of growth, R_T/T, of a well designed on-line learning algorithm typically diminishes, which implies that, for the infinite horizon

problem, the profit achieved per unit time without knowing the demand function matches that when the demand function is known. Therefore, a relevant performance metric is the growth rate of regret R_T vs. T.x

1.4.1 Summary of results

The basic problem setting involves two players: a retailer (an electricity distributor or aggregator) who offers its customer day-ahead hourly dynamic prices and its customers with price responsive demands. We focus on the case when the customer demands are elastic and can be described by a random affine model, which arises naturally for thermal control applications.

The main result of our work is twofold. First, under the DAHP mechanism, we propose a simple online learning algorithm, referred to as piecewise linear stochastic approximation (PWLSA), that has the logarithmic rate of growth in regret, i.e. $R_T(T) = \Theta(\log T)$.

On the other hand, we show that no other on-line learning algorithm can have the rate slower than that of PWLSA. Thus PWLSA is order optimal. To achieve the optimal rate of learning, we deviate the standard on-line learning approach by first analyzing the mechanism of the two-settlement wholesale electricity market and calculate the retail surplus of the retailer as a wholesale market participant in a simple set-up. The result shows that the retailer's loss of surplus is proportional to the 2-norm deviation of the real-time consumption from the day-ahead schedule.

To demonstrate the learning performance, we also conduct simulations to

compare PWLSA with the Greedy Method based on the actual data. In both cases with static and dynamically changing parameters of the demand model, PWLSA outperformed the greedy method and converged fast towards the optimal price.

1.4.2 Related work

The problem of dynamic pricing for demand response assuming known demand functions has been extensively studied. See, for example, [12, 11, 25], which adopted a similar pricing scheme as considered here and [15, 17, 71] for more general settings. A precursor of the work presented here is [29] where a parametric form of demand function was obtained. In [30], the tradeoff between retail profit and consumer surplus was characterized under a Stackelberg formulation with known demand functions.

The general problem of online learning for dynamic pricing has been studied extensively in multiple communities. This problem can be formulated as a multi-armed bandit (MAB) problem by treating each possible price as an arm. When the price can only take finite possible values, the problem becomes the classic MAB for which Lai and Robbins showed that the optimal regret growth rate is $\Theta(\log T)$ when the arms generate independent reward [42]. When the price takes value from an uncountable set, the dynamic pricing problem is an example of the so-called continuum-armed bandit introduced by Agrawal in [6] where the arms form a compact subset of $\mathcal{R}$. An online learning policy with regret order of $O(T^{3/4})$ was proposed in [6] for any reward function satisfying Lipschitz continuity. Further development on the continuum-armed

bandit under various assumptions of the unknown reward function can be found in [36, 8, 18]. The reason that our proposed PWLSA achieves a much better regret order ($\Theta(\log T)$) than in the case of a general continuum-armed bandit is due to the specific linearly parameterized demand which leads to a specific quadratic cost/reward function. A similar message can be found in [37, 58, 13, 24, 72] where different regret orders were shown to be achievable under different classes of demand models for dynamic pricing.

The problem considered here deals with linearly parameterized demand functions, thanks to the closed-form characterization of the optimized demand function for thermal dynamic load obtained in [30]. Our proposed learning approach is rooted from a stochastic approximation problem originally formulated by Lai and Robbins [40, 41] where the authors considered a form of optimal control problem when the model contains unknown parameters and the cost of control is explicitly modeled. For scaler models, Lai and Robbins showed in [40, 41] that the cumulative regret (if translated from our definition) of a simple linear stochastic approximation scheme grows at the rate of $O(\log T)$. However, it is not clear whether such growth rate is the lowest possible. Our result provides a generalization to the vector case with a lower bound for general policies. In addition, our approach also allows the consumers to have variable demand levels whereas the algorithm presented in [40, 41] only allows a single constant demand target.

Also related is the work of Bertsimas and Perakis [10] who tackled the problem as a dynamic program with incomplete state information. The authors showed in numerical simulations that considerable gain can be realized over the myopic policy where the price in the next stage is based on the least squares

estimate of the model parameter. When the parameters are assumed to be random, Lobo and Boyd considered the same problem under a Bayesian setting [47] and proposed a randomized policy via a dithering mechanism. In both cases, the rate of learning is not characterized.

Machine learning techniques have been applied to pricing problems in electricity markets, although there seems to be limited literature on discovering real-time price with unknown demand functions at the retail level. While such problems can be viewed as part of the general learning problem discussed above, the nature of electricity market and electricity demand impose special constraints. When the market has multiple strategic generators, Garcia et al. proposed an online learning algorithm which converges to the Markov perfect equilibria [21]. A related learning problem of bidding strategy of a retailer in the wholesale market when the supply functions of the generators are unknown has been studied. See [57, 56, 55] where Q-learning techniques have been applied. Some other research focuses on developing learning methods for optimal demand response. See [65] for index policy by formulating the demand control as a restless bandit problem, and [51] for a reinforcement learning solution to a partially observable Markov decision process (MDP) problem.

CHAPTER 2

EFFECT OF DATA QUALITY ON REAL-TIME WHOLESALE PRICE

2.1 Structures of Real-Time LMP

We present first a model for the computation of real-time locational marginal price (LMP). While ISOs have somewhat different methods of computing real-time LMP, they share the same two-settlement architecture and similar ways of using real-time measurements. In the following, we will use a simplified ex-post real-time market model, adopted by PJM, ISO New England, and other ISOs [52, 73]. We view this model as a convenient mathematical abstraction that captures the essential components of the real-time LMP calculation. For this reason, our results should be interpreted within the specified setup. Our purpose is not to include all details; we aim to capture the essential features.

In real-time, in order to monitor and operate the system, the control center will calculate the estimated system conditions (including bus voltages, branch flows, generation, and demand) based on real-time measurements. We call a branch congested if the estimated flow is larger than or equal to the security limit. The congestion pattern is defined as the set of all congested lines, denoted as $\hat{\mathcal{C}}$. Note that we use hat (*e.g.,* $\hat{\mathcal{C}}$) to denote quantities or sets that are estimated based on real-time measurements. Details of state estimation and bad data detection are discussed in Section 2.2.3.

One important usage of state estimation is calculating the real-time LMP. Given the estimated congestion pattern $\hat{\mathcal{C}}$, the following linear program is solved to find the incremental OPF dispatch and associated real-time LMP,

$\hat{\lambda} = (\hat{\lambda}_i)$ [52]:

$$\begin{aligned} \text{minimize} \quad & \sum c_i^{\mathrm{G}} \Delta p_i - \sum c_j^{\mathrm{L}} \Delta d_j \\ \text{subjcet to} \quad & \sum \Delta p_i = \sum \Delta d_j \\ & \Delta p_i^{\min} \le \Delta p_i \le \Delta p_i^{\max} \\ & \Delta d_j^{\min} \le \Delta d_j \le \Delta d_j^{\max} \\ & \textstyle\sum_i A_{ki} \Delta p_i - \sum_j A_{kj} \Delta d_j \le 0, \text{for all } k \in \hat{\mathcal{C}}, \end{aligned} \tag{2.1}$$

where $\Delta d = (\Delta d_j)$ is the vector of incremental dispatchable load, $\Delta p = (\Delta p_i)$ the vector of incremental generation dispatch, $c^{\mathrm{G}} = (c_i^{\mathrm{G}})$ and $c^{\mathrm{L}} = (c_j^{\mathrm{L}})$ the corresponding real-time marginal cost of generations and dispatchable loads, $\Delta p_i^{\min}$ and $\Delta p_i^{\max}$ the lower and upper bounds for incremental generation dispatch, $\Delta d_i^{\min}$ and $\Delta d_i^{\max}$ the lower and upper bounds for incremental dispatchable load, and A_{ki} the sensitivity of branch flow on branch k with respect to the power injection at bus i.

The real-time LMP at bus i is defined as the overall cost increase when one unit of extra load is added at bus i, which is calculated as

$$\hat{\lambda}_i = \eta - \sum_{k \in \hat{\mathcal{C}}} A_{ki} \mu_k. \tag{2.2}$$

where η is the dual variable for the load-generation equality constraint, and μ_k is the dual variable corresponding to the line flow constraint in (2.1).

Note that in practice, the control center may use the ex-ante congestion pattern, which is obtained by running a 5 minute ahead security-constrained economic dispatch with the state estimation results and the forecasted loads (for the next five-minute interval) and choosing the lines congested at the dispatch solution [52, 73]. However, to avoid the complication due to ex-ante dispatch calculation, we assume that real-time pricing employs the estimated congestion pattern $\hat{\mathcal{C}}$ obtained from state estimation results. By doing so, we attempt to find

direct relations among bad data, the state estimate, and real-time LMPs. Notice that once the congestion pattern $\hat{\mathcal{C}}$ is determined, the whole incremental OPF problem (2.1) no longer depends on the measurement data.

Under the DC model, the power system state, x, is defined as the vector of voltage phases, except the phase on the reference bus. The power flow vector f is a function of the system state x,

$$f = Fx, \tag{2.3}$$

where F is the sensitivity matrix of branch flows with respect to the system state.

Assume the system has $n+1$ buses. Then, $x \in \mathcal{X} = [-\pi, \pi]^n$, where $\mathcal{X}$ represents the state space. Any system state corresponds to a unique point in $\mathcal{X}$. From (2.3), the branch flow f is determined by the system state x. Comparing the flows with the flow limits, we obtain the congestion pattern associated with this state. Hence, each point in the state space corresponds to a particular congestion pattern.

We note that the above expression in (2.2) appears earlier in [68] where the role of congestion state in LMP computation was discussed. Here, our objective is to make explicit the connection between data and LMP. We therefore need a linkage between data and congestion. To this end, we note that the power system state, the congestion state, and LMP form a Markov chain, which led to a geometric characterization of LMP on the power system state space, as shown in the following theorem.

Theorem 1 (Price Partition of the State Space) *Assume that the LMP exists for every possible congestion pattern*[*]*. Then, the state space $\mathcal{X}$ is partitioned into a set*

[*]This is equivalent to assuming that the derivative of the optimal value of (2.1) with respect

of polytopes $\{\mathcal{X}_i\}$ where the interior of each $\mathcal{X}_i$ is associated with a unique congestion pattern $\mathcal{C}_i$ and a real-time LMP vector. Each boundary hyperplane of $\mathcal{X}_i$ is defined by a single transmission line.

Proof 1 *For a particular congestion pattern $\mathcal{C}$ defined by a set of congested lines, the set of states that gives $\mathcal{C}$ is given by*

$$\mathcal{X}_i \stackrel{\Delta}{=} \{x : F_{i\cdot} x \geq T_i^{max} \ \forall i \in \mathcal{C}, F_{j\cdot} x < T_j^{max} \ \forall j \notin \mathcal{C}\},$$

where $F_{i\cdot}$ is the ith row of F (see (2.3)), and T_j^{max} the flow limit on branch j. Since $\mathcal{X}_i$ is defined by the intersection of a set of half spaces, it is a polytope.

Given an estimated congestion pattern $\hat{\mathcal{C}}$, the envelop theorem [49] implies that for any optimal primal solution and dual solution of (2.1) that satisfy the KKT conditions, (2.2) always gives the derivative of the optimal objective value with respect to the demand at each bus, which we assume exists, i.e., each congestion pattern is associated with a unique real-time LMP vector λ. Hence, all states with the same congestion pattern share the same real-time LMP, which means each polytope $\mathcal{X}_i$ in $\mathcal{X}$ corresponds to a unique real-time LMP vector.

Theorem 1 characterizes succinctly the relationship between the system state and LMP. As illustrated in Fig. 1.1(a), if bad data are to alter the LMP in real-time, the size of the bad data has to be sufficiently large so that the state estimate at the control center is moved to a different price region from the true system state.

On the other hand, if some lines are erroneously removed from or added to the correct topology, as illustrated in Fig. 1.1(b), it affects the LMP calculation

to demand at each bus exists

in three ways[†]. First, the state estimate is perturbed since the control center employs an incorrect topology in state estimation. Secondly, the price partition of the state space changes due to the errors in topology information. Third, the shift matrix A in (2.1), which is a function of topology, changes thereby altering prices attached to each price region.

2.2 Data Model and State Estimation

2.2.1 Meter data

In order to monitor the system, various meter measurements are collected in real time, such as power injections, branch flows, voltage magnitudes, and phasors, denoted by a vector $z \in \mathbb{R}^m$. [‡] If there exists bad data a among the measurements, the measurements with bad data, denoted by z_a, can be expressed as a function of the system states x,

$$z_a = z + a = h(x) + w + a, \quad a \in \mathcal{A}, \tag{2.4}$$

where w represents the random measurement noise.

We make a distinction here between the measurement noise and bad data; the former accounts for random noise independently distributed across all meters whereas the latter represents the perturbation caused by bad or malicious

[†]In addition to these, the change in topology will affect contingency analysis. Such effect will appear as changes in contingency constraints in real-time LMP calculation (2.1) [52]. However, dealing with contingency constraints will significantly complicate our analysis and possibly obscure the more direct link between bad data and real-time LMP. Hence, we consider only line congestion constraints in (2.1).

[‡]Notice here both conventional measurements and PMU measurements can be incorporated. Although PMU data seem to have more direct impact on state estimation and real-time LMP calculation, we won't differentiate the types of measurements in the following discussion.

data. We assume no specific pattern for bad data except that they do not happen everywhere. We assume that bad data can only happen in a subset of the measurements, $\mathcal{S}$. We call $\mathcal{S}$ as set of suspectable meters, which means the meter readings with in $\mathcal{S}$ may subject to corruption. If the cardinality of $\mathcal{S}$ is k, the feasible set of bad data a is a k-dimensional subspace, denoted as $\mathcal{A} = \{a : a_i = 0 \text{ for all } i \notin \mathcal{S}\}$.

We will consider three bad data models with increasing power of affecting state estimates.

M1. *State independent bad data*: This type of bad data is independent of real-time measurements. Such bad data may be the replacement of missing measurements.

M2. *Partially adaptive bad data*: This type of bad data may arise from the so-called man in the middle (MiM) attack where an adversary intercepts the meter data and alter the data based on what he has observed. Such bad data can adapt to the system operating state.

M3. *Fully adaptive bad data*: This is the most powerful type of bad data, constructed based on the actual measurement $z = h(x) + w$.

Note that M3 is in general not realistic. Our purpose of considering this model is to use it as a conservative proxy to obtain performance bounds for the impact of worst case data.

We assume herein a DC model in which the measurement function $h(\cdot)$ in (2.4) is linear. Specifically,

$$z_a = Hx + w + a, \quad a \in \mathcal{A}, \tag{2.5}$$

where H is the measurement matrix. Such a DC model, while widely used in the literature, may only be a crude approximation of the real power system. By making such a simplifying assumption and acknowledging its weaknesses, we hope to obtain tractable solutions in searching for worst case scenarios. It is important to note that, although the worst case scenarios are derived from the DC model, we carry out simulations using the actual nonlinear system model.

2.2.2 Topology data

Topology data are represented by a binary vector $s \in \{0,1\}^l$, where each entry of s represents the state of a line breaker (0 for open and 1 for closed). The bad topology data is modeled as

$$s_b = s + b \pmod 2, \quad b \in \mathcal{B}, \tag{2.6}$$

where $\mathcal{B} \subset \{0,1\}^l$ is the set of possible bad data. When bad data are present, the topology processor will generate the topology estimate corresponding to s_b, and this incorrect topology estimate will be passed to the following operations unless detected by the bad data detector.

2.2.3 State Estimation

We assume that the control center employs the standard weighted least squares (WLS) state estimator. Under DC model,

$$\hat{x} = \arg\min_x (z - Hx)^{\mathrm{T}} R^{-1} (z - Hx) = Kz, \tag{2.7}$$

where R is the covariance matrix of measurement noise w, and $K \triangleq (H^{\mathrm{T}} R^{-1} H)^{-1} H^{\mathrm{T}} R^{-1}$.

If the noise w is Gaussian, the WLS estimator is also the maximum likelihood estimate (MLE) of state x. By the invariant property of MLE, from (2.3), the maximum likelihood estimate of the branch flows is calculated as

$$\hat{f} = F\hat{x} = FKz. \tag{2.8}$$

The congestion pattern used in real-time LMP calculation (2.1) is directly from state estimation and consists of all the estimated branch flows which are larger than or equal to the branch flow limits, *i.e.*,

$$\hat{\mathcal{C}} = \{j : \hat{f}_j \geq T_j^{\max}\}, \tag{2.9}$$

where $T_j^{\max}$ is the flow limit on branch j.

In the presence of bad meter data a, the meter measurements collected by control center is actually $z_a = Hx + w + a$. By using z_a, the WLS state estimate is

$$\hat{x}_a = Kz_a = \hat{x}^* + Ka, \tag{2.10}$$

where $\hat{x}^* = Kz$ is the "correct" state estimate without the presence of the bad data (*i.e.*, $a = 0$).

Eq. (2.10) shows that the effect of bad data on state estimation is linear. However, because a is confined in a k-dimensional subspace $\mathcal{A}$, the perturbation on the actual system state is limited to a certain direction.

When bad data exist both in meter and topology data, the control center uses a wrong measurement matrix $\bar{H}$, corresponding to the altered topology data, and the altered meter data z_a. Then, the WLS state estimate becomes

$$\hat{x}_a = \bar{K}z_a = \bar{K}z + \bar{K}a, \tag{2.11}$$

where $\bar{K} \triangleq (\bar{H}^{\mathrm{T}} R^{-1} \bar{H})^{-1} \bar{H}^{\mathrm{T}} R^{-1}$. Note that unlike the linear effect of bad meter data, bad topology data affects the state estimate by altering the measurement matrix H to $\bar{H}$.

2.2.4 Bad Data Detection

The control center uses bad data detection to minimize the impact of bad data. Here, we assume a standard bad data detection used in practice, the $J(\hat{x})$-detector in [23]. In particular, the $J(\hat{x})$-detector performs the test on the residue error, $r \triangleq z - H\hat{x}$, based on the state estimate $\hat{x}$. From the WLS state estimate (2.7), we have

$$r = \left(I - H(H^{\mathrm{T}} R^{-1} H)^{-1} H^{\mathrm{T}} R^{-1}\right) z = Uz. \tag{2.12}$$

where $U \stackrel{\Delta}{=} (I - H(H^{\mathrm{T}} R^{-1} H)^{-1} H^{\mathrm{T}} R^{-1})$

The $J(\hat{x})$-detector is a threshold detector defined by

$$r^{\mathrm{T}} R^{-1} r = z^{\mathrm{T}} W z \underset{\text{good data}}{\overset{\text{bad data}}{\gtrless}} \tau, \tag{2.13}$$

where τ is the threshold calculated from a prescribed false alarm probability, and $W \stackrel{\Delta}{=} U^{\mathrm{T}} R^{-1} U$. When the measurement data fail to pass the bad data test, the control center declares the existence of bad data and takes corresponding actions to identify and remove the bad data.

In the following, we are interested in those cases when bad data are present while the $J(\hat{x})$-detector fails to detect them.

2.3 Impact of Bad Data on LMP

In this section, we examine the impact of bad data on LMP, assuming that the topology estimate of the network is correct.

One thing to notice is that in searching for the "worst" case, we take the perspective of the control center, not that of the attacker. In particular, we look for the worst congestion pattern for the LMP computation, even if this particular congestion pattern is difficult for the attacker to discover. So the focus here is not how easy it is for an attacker to find a locally worst congestion pattern; it is how much such a congestion pattern affects the LMP.

2.3.1 Average Relative Price Perturbation

In order to quantify the effect of bad data on real-time price, we need to first define the metric to measure the effect. We define the relative price perturbation (RPP) as the expected percentage price perturbation caused by bad data. Given that LMP varies at different buses, RPP also varies at different locations.

Let z_a be the data received at the control center and $\lambda_i(z_a)$ the LMP at bus i. The RPP at bus i is a function of bad data a, given by

$$\mathrm{RPP}_i(a) = \mathbb{E}\left(\left|\frac{\lambda_i(z_a) - \lambda_i(z)}{\lambda_i(z)}\right|\right), \tag{2.14}$$

where the expectation is over random state and measurement noise.

To measure the system-wide price perturbation, we define the average rela-

tive price perturbation (ARPP) by

$$\mathrm{ARPP}(a) = \frac{1}{n+1} \sum_i \mathrm{RPP}_i(a), \tag{2.15}$$

where $n + 1$ is the number of buses in the system.

The worst case analysis to be followed can be used for other metrics (e.g., price increase ratios or price decrease ratios, which are closely related to the market participants' gain or loss). Similar results can be showed following the same strategies. However, the comparison among different metrics is beyond the scope of our work.

2.3.2 Worst ARPP under State Independent Bad Data Model

First, we consider the state independent bad data model (M1) given in Section 2.2.1. In this model, the bad data are independent of real-time measurements.

In constructing the state independent worst data, it is useful to incorporate prior information about the state. To this end, we assume that system state follows a Gaussian distribution with mean x_0, covariance matrix Σ_x. Typically, we choose x_0 as the day-ahead dispatch since the nominal system state in real-time varies around its day-ahead projection.

In the presence of bad data a, the expected state estimate and branch flow estimate on branch i are given by

$$\mathbb{E}[\hat{x}] = x_0 + Ka. \tag{2.16}$$

$$\mathbb{E}[f_i] = F_{i\cdot}\mathbb{E}[\hat{x}] = F_{i\cdot}x_0 + F_{i\cdot}Ka, \tag{2.17}$$

where $F_{i\cdot}$ is the corresponding row of branch i in F.

Our strategy is to make this expected state estimate into the region with the largest price perturbation among all the possible regions, $\hat{\mathcal{C}}^*$. From (2.9), this means making all the expected branch flows satisfy the boundary condition of $\hat{\mathcal{C}}^*$,

$$\begin{aligned} \mathbb{E}[f_i] &\geq T_i^{\max} \quad \text{for } i \in \hat{\mathcal{C}}^* \\ \mathbb{E}[f_i] &\leq T_j^{\max} \quad \text{for } j \notin \hat{\mathcal{C}}^*. \end{aligned} \tag{2.18}$$

However, due to the uncertainty (from both system state x and measurement noise w), the actual estimated state after attack, $\hat{x}$, may be different from $\mathbb{E}[\hat{x}]$. Therefore, we want to make $\mathbb{E}[\hat{x}]$ at the "center" of the desired price region, *i.e.*, maximizing the shortest distance from $\mathbb{E}[\hat{x}]$ to the boundaries of the polytope price regions while still holding the boundary constraints. The shortest distance can be calculated as

$$\beta = \min\{\tilde{\beta} : |\mathbb{E}[f_i] - T^{\max}| \geq \tilde{\beta} \text{ for all i}\}. \tag{2.19}$$

However, the existence of bad data detector prevents the bad data vector a from being arbitrarily large. According to (2.12), the weighted squared residue with a is

$$r^{\mathrm{T}} R^{-1} r = z_a^{\mathrm{T}} W z_a = (w + a)^{\mathrm{T}} W (w + a). \tag{2.20}$$

since $WHx = 0$

Heuristically, since w has zero mean, the term $a^{\mathrm{T}} W a$ can be used to quantify the effect of data perturbation on estimation residue. Then we use $a^{\mathrm{T}} W a \leq \epsilon$ to control the detection probability in the following optimization.

Therefore, for a specific congestion pattern $\hat{\mathcal{C}}$, the adversary will solve the following optimization problem to move the state estimate to the "center" of

the price region $\hat{\mathcal{C}}$ and keeping the detection probability low.

$$\begin{aligned} \max_{a\in\mathcal{A},\tilde{\beta}\geq 0} \quad & \tilde{\beta} \\ \text{subject to} \quad & \mathbb{E}[f_i] - \tilde{\beta} \geq T_i^{\max}, i \in \hat{\mathcal{C}} \\ & \mathbb{E}[f_i] + \tilde{\beta} < T_j^{\max}, j \notin \hat{\mathcal{C}} \\ & a^{\mathrm{T}} W a \leq \epsilon, \end{aligned} \tag{2.21}$$

which is a convex program that can be solved easily in practice. We call a region $\hat{\mathcal{C}}$ feasible if it makes problem (2.21) feasible.

Among all the feasible congestion patterns, the worst region $\hat{\mathcal{C}}^*$ is chosen as the one giving the largest ARPP.

$$\hat{\mathcal{C}}^* = \arg \max_{\hat{\mathcal{C}}\in\Gamma} |\tilde{\lambda}_i - \lambda_i(\hat{\mathcal{C}})|, \tag{2.22}$$

where $\tilde{\lambda}_i$ is the LMP at bus i if the x_0 is the system state, and Γ the set of all the feasible congestion patterns. Hence, the worst case constant bad data vector is the solution to optimization problem (2.21) by setting the congestion pattern as $\hat{\mathcal{C}}^*$.

2.3.3 Worst ARPP under Partially Adaptive Bad Data

For bad data model M2, only part of the measurement values in real-time are known to the adversary, denoted as z_{o}. The adversary has to first make an estimation of the system state from the observation and prior distribution, then make the attack decision based on the estimation result.

Without the presence of bad data vector, *i.e.*, $a = 0$, the system equation (2.5) gives

$$z_{\mathrm{o}} = H_{\mathrm{o}} x + w_{\mathrm{o}}, \tag{2.23}$$

where H_o is the rows of H corresponding to the observed measurements and w_o the corresponding part in the measurement noise w.

The minimum mean square error (MMSE) estimate of x given z_o is given by the conditional mean

$$\mathbb{E}(x|z_o) = x_0 + \Sigma_x H_o^T (H_o \Sigma_x H_o^T)^{-1}(z_o - H_o x_0). \tag{2.24}$$

Then, the flow estimate on branch i after attack is

$$\mathbb{E}[f_i|z_o] = F_i.\mathbb{E}[\hat{x}|z_o]. \tag{2.25}$$

Still, we want to move the estimation of state to the "center". On the other hand, the expected measurement value $\mathbb{E}[z_a|z_o] = H\mathbb{E}[\hat{z}|z_o] + a$. Again, we need a pre-designed parameter ϵ to control the detection probability. Therefore, the solution to the following optimization problem is the best attack given congestion pattern $\mathcal{A}$

$$\begin{aligned} \max_{a\in\mathcal{A},\tilde{\beta}\geq 0} \quad & \tilde{\beta} \\ \text{subject to} \quad & \mathbb{E}[f_i|z_o] - \tilde{\beta} \geq T_i^{\max}, i \in \hat{\mathcal{C}} \\ & \mathbb{E}[f_i|z_o] + \tilde{\beta} < T_j^{\max}, j \notin \hat{\mathcal{C}} \\ & (H\mathbb{E}[z_a|z_o]^T) W (H\mathbb{E}[z_a|z_o]) \leq \epsilon. \end{aligned} \tag{2.26}$$

This problem is also a convex optimization problem, which can be easily solved. Among all the $\hat{\mathcal{C}}$'s which make the above problem feasible, we choose the one with the largest price perturbation, denoted as $\hat{\mathcal{C}}^*$. The solution to problem (2.26) with $\hat{\mathcal{C}}^*$ as the congestion pattern is the worst bad data vector.

2.3.4 Worst ARPP under Fully Adaptive Bad Data

Finally, we consider the bad data model M3, in which the whole set of measurements z is known to the adversary. The worst bad data vector depends on the value of z. Different from the previous two models, with bad data vector a, the estimated state is deterministic without uncertainty. In particular

$$\hat{x} = Kz + Ka. \tag{2.27}$$

And the estimated flow on branch i after attack is also deterministic

$$\hat{f}_i = F_{i\cdot}\hat{x} = F_{i\cdot}Kz + F_{i\cdot}Ka. \tag{2.28}$$

Similar to the previous two models, congestion pattern is called feasible if there exists some bad data vector a to make the following conditions satisfied:

$$\begin{aligned} &\hat{f}_i \geq T_i^{\max}, i \in \hat{\mathcal{C}} \\ &\hat{f}_i < T_j^{\max}, j \notin \hat{\mathcal{C}} \\ &(z+a)^{\mathrm{T}} W (z+a) \leq \tau, \quad a \in \mathcal{A}. \end{aligned} \tag{2.29}$$

Among all the feasible congestion patterns, we choose the one with the largest price perturbation, $\hat{\mathcal{C}}^*$. Any bad data vector a satisfying condition (2.29) can serve as the worst fully adaptive bad data.

2.3.5 A Greedy Heuristic

The strategies presented above are based on the exhaustive search over all possible congestion patterns. Such approaches are not scalable for large networks with a large number of possible congestion patterns. We now present a greedy

heuristic approach aimed at reducing computation cost. In particular, we develop a gradient like algorithm that searches among a set of likely congestion patterns.

First, we restrict ourselves to the set of lines that are close to their respective flow limits and look for bad data that will affect the congestion pattern. The intuition is that it is unlikely that bad data can drive the system state sufficiently far without being detected by the bad data detector. In practice, the cardinality of such a set is usually very small compared with the systems size.

Second, we search for the worst data locally by changing one line in the congestion pattern at a time. Specifically, suppose that a congestion pattern is the current candidate for the worst data. Given a set of candidate lines that are prone to congestions, we search locally by flipping one line at a time from the congested state to the un-congested state and vice versa. If no improvement can be made, the algorithm stops. Otherwise, the algorithm updates the current "worst congestion pattern" and continue. The effectiveness of this greedy heuristic is tested in Section 2.5.3.

2.4 Bad Topology Data on LMP

So far, we have considered bad data in the analog measurements. In this section, we include the bad topology data, and describe another bad data model.

We represent the network topology by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where each $i \in \mathcal{V}$ denotes a bus and each $(i, j) \in \mathcal{E}$ denotes a connected transmission line. For each physical transmission line (*e.g.,* a physical line between i and j),

we assign an arbitrary direction (*e.g.,* (i, j)) for the line, and (i, j) is in $\mathcal{E}$ if and only if bus i and bus j are connected.

Bad data may appear in both analog measurements and digital (*e.g.,* breaker status) data, as described in Section 2.2.1 and Section 2.2.2.

$$\begin{aligned} z_a &= z + a = (Hx + w) + a, \quad a \in \mathcal{A}, \\ s_b &= s + b \pmod 2, \quad b \in \mathcal{B}. \end{aligned} \tag{2.30}$$

As in Section 2.3, we employ the adversary model to describe the worst case. The adversary alters s to s_b by adding b from the set of feasible attack vectors $\mathcal{B} \subset \{0, 1\}^l$ such that the topology processor produces the "target" topology $\bar{\mathcal{G}}$ as the topology estimate. In addition, the adversary modifies z by adding $a \in \mathcal{A}$ such that z_a looks consistent with $\bar{\mathcal{G}}$.

In this section, we focus on the worst case when the adversary is able to alter the network topology without changing the state estimate[§]. We also require that such bad data are generated by an adversary causing undetectable topology change, *i.e.,* the bad data escape the system bad data detection. For the worst case analysis, we will maximize the LMP perturbation among the attacks within this specific class. Even though this approach is suboptimal, the simulation results in Section 2.5 demonstrate that the resulting LMP perturbation is much greater than the worst case of the bad meter data.

Suppose the adversary wants to mislead the control center with the target topology $\bar{\mathcal{G}} = (\mathcal{V}, \bar{\mathcal{E}})$, a topology obtained by removing[¶] a set of transmission

[§]In general, the adversary can design the worst data to affect both the state estimate and network topology. It is, however, much more difficult to make such attack undetectable.

[¶]Line addition by the adversary is also possible [33]. However, compared to line removal attacks, line addition attacks require the adversary to observe a much larger set of meter measurements to design undetectable attacks. In addition, the number of necessary modifications

lines $\mathcal{E}_\Delta$ in $\mathcal{G}$ (*i.e.*, $\bar{\mathcal{E}} = \mathcal{E} \setminus \mathcal{E}_\Delta$). We assume that the system with $\bar{\mathcal{G}}$ is observable: *i.e.*, the corresponding measurement matrix $\bar{H}$ has full column rank[‖].

The adversarial data modification aimed at perturbing the topology estimate at the control center was studied in [34]. Suppose that the adversary changes the breaker status such that the target topology $\bar{\mathcal{G}} = (\mathcal{V}, \bar{\mathcal{E}})$ is observed at the control center. Simultaneously, if the adversary introduces bad data $a = \bar{H}x - Hx$, then

$$z_a = Hx + a + w = \bar{H}x + w, \tag{2.31}$$

which means that the meter data received at the control center are completely consistent with the model generated from $\bar{\mathcal{G}}$. Thus, any bad data detector will not be effective.

It is of course not obvious how to produce the bad data a, especially when the adversary can only modify a limited number of measurements, and it may not have access to the entire state vector x. Fortunately, it turns out that a can be generated by observing only a few entries in z without requiring global information (such as the state vector x) [34].

A key observation is that Hx and $\bar{H}x$ differ only in a few entries corresponding to the modified topology (lines in $\mathcal{E}_\Delta$) as illustrated in Fig. 2.1. Consider first the noiseless case. Let z_{ij} denote the entry of z corresponding to the flow measurement from i to j. As hinted from Fig. 2.1, it can be easily seen that $\bar{H}x - Hx$ has the following sparse structure [34]:

$$\bar{H}x - Hx = -\sum_{(i,j)\in\mathcal{E}_\Delta} \alpha_{ij} m_{(i,j)}, \tag{2.32}$$

in breaker data is also much larger: to make a line appear to be connected, the adversary should make all the breakers on the line appear to be closed. Please see [34] for the detail.

[‖]Without observability, the system may not proceed to state estimation and real-time pricing. Hence, for the adversary to affect pricing, the system with the target topology has to be observable.

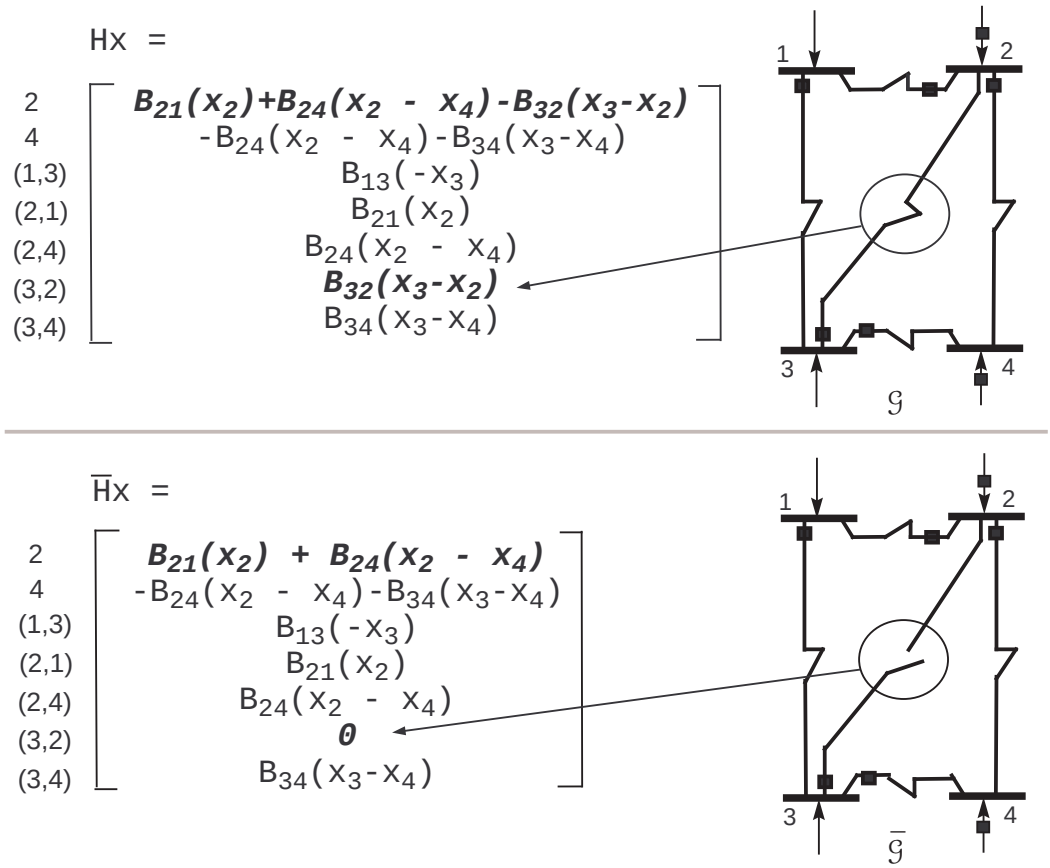


Figure 2.1: Hx and $\bar{H}x$: Each row is marked by the corresponding meter (i for injection at i and (i, j) for flow from i to j).

where $\alpha_{ij} \in \mathbb{R}$ denotes the line flow from i to j when the line is connected and the system state is x, and $m_{(i,j)}$ is the column of the measurement-to-branch incidence matrix, that corresponds to (i, j): *i.e.*, $m_{(i,j)}$ is an m-dimensional vector with 1 at the entries corresponding to the flow from i to j and the injection at i, and -1 at the entries for the flow from j to i and the injection at j, and 0 at all other entries. Absence of noise implies that $z_{ij} = \alpha_{ij}$, which leads to

$$\bar{H}x - Hx = -\sum_{(i,j)\in\mathcal{E}_\Delta} z_{ij} m_{(i,j)}. \tag{2.33}$$

With (2.33) in mind, one can see that setting $a = \bar{H}x - Hx$ and adding a to z is equivalent to the following simple procedure: as described in Fig. 2.2, for each

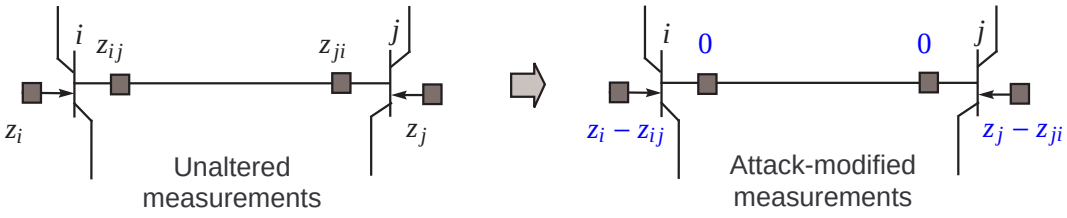


Figure 2.2: The attack modifies local measurements around the line (i, j) in $\mathcal{E}_\Delta$.

(i, j) in $\mathcal{E}_\Delta$,

1. Subtract z_{ij} and z_{ji} from z_i and z_j respectively.
2. Set z_{ij} and z_{ji} to be 0.

where z_i is the entry of z corresponding to the injection measurement at bus i.

When measurement noise is present (*i.e.*, $z = Hx + w$), the idea of the attack is still the same: to make a approximate $\bar{H}x - Hx$ so that z_a is close to $\bar{H}x + w$. Since $z_{ij} = \alpha_{ij} + w_{ij}$, z_{ij} is an unbiased estimate of α_{ij} for each $(i, j) \in \mathcal{E}_\Delta$, and this implies that $-\sum_{(i,j)\in\mathcal{E}_\Delta} z_{ij} m_{(i,j)}$ is an unbiased estimate of $-\sum_{(i,j)\in\mathcal{E}_\Delta} \alpha_{ij} m_{(i,j)} = \bar{H}x - Hx$. Hence, we set a to be $-\sum_{(i,j)\in\mathcal{E}_\Delta} z_{ij} m_{(i,j)}$, the same as in the noiseless setting, and the attack is executed by the same steps as above.

For launching this attack to modify the topology estimate from $\mathcal{G}$ to $\bar{\mathcal{G}}$, the adversary should be able to (i) set b such that the topology processor produces $\bar{\mathcal{G}}$ instead of $\mathcal{G}$ and (ii) observe and modify z_{ij}, z_{ji}, z_i, and z_j for all $(i, j) \in \mathcal{E}_\Delta$. The attack is feasible if and only if $\mathcal{A}$ and $\mathcal{B}$ contain the corresponding attack vectors.

To find the worst case LMP perturbation due to undetectable, state-preserving attacks, let $\mathcal{F}$ denote the set of feasible $\bar{\mathcal{G}}$s, for which the attack can be

launched with $\mathcal{A}$ and $\mathcal{B}$. Among the feasible targets in $\mathcal{F}$, we consider the best target topology that results in the maximum perturbation in real-time LMPs. If ARPP is used as a metric, the best target is chosen as

$$\bar{\mathcal{G}}^*[z] = \arg \max_{\bar{\mathcal{G}} \in \mathcal{F}} \sum_i \left| \frac{\lambda_i(z; \bar{\mathcal{G}}) - \lambda_i(z; \mathcal{G})}{\lambda_i(z; \mathcal{G})} \right|. \tag{2.34}$$

where $\lambda_i(z; \bar{\mathcal{G}})$ denotes the real-time LMP at bus i when the attack with the target $\bar{\mathcal{G}}$ is launched on z, and $\lambda_i(z; \mathcal{G})$ is the real-time LMP under no attack.

2.5 Numerical Results

In this section, we demonstrate the impact of bad data on real-time LMPs with the numerical simulations on IEEE-14 and IEEE-118 systems. We conducted simulations in two different settings: the linear model with the DC state estimator and the nonlinear model with the AC state estimator. The former is usually employed in the literature for the ease of analysis whereas the latter represents the practical state estimator used in the real-world power system. In all simulations, the meter measurements consist of real power injections at all buses and real power flows (both directions) at all branches.

2.5.1 Linear model with DC state estimation

We first present the simulation results for the linear model with the DC state estimator. We modeled bus voltage magnitudes and phases as Gaussian random variables with the means equal to the day-ahead dispatched values and small standard deviations. In each Monte Carlo run, we generated a state realization from the statistical model, and the meter measurements were created by

the DC model with Gaussian measurement noise. Once the measurements were created, bad data were added in the manners discussed in Section 2.3 and Section 2.4. With the corrupted measurements, the control center executed the DC state estimation and the bad data test with the false alarm probability constraint 0.1. If the data passed the bad data test, real-time LMPs were evaluated based on the state estimation results. For IEEE-14 and IEEE-118 system, the network parameters[**] are available in [4].

We used the number of meter data to be modified by the adversary as the metric for the attack effort. For the 14 bus system, in each Monte Carlo run, we randomly chose two lines, and the adversary was able to modify all the line flow meters on the lines and injection meters located at the ends of the lines. For the 118 bus system, we randomly chose three lines, and the adversary had control over the associated line and injection meters. Both state and topology attacks were set to control the same number of meter data[††] so that we can fairly compare their impacts on real-time LMPs. As for the meter data attack, we only considered the lines that are close to their flow limits (estimated flows under M1 and M2, or actual flows under M3) as candidates for congestion pattern search. The threshold is chosen as 10MW in our simulation.

[**]In addition to the network parameters given in [4], we used the following line limit and real-time offer parameters. In the IEEE-14 simulation, the generators at the buses 1, 2, 3, 6, and 8 had capacities 330, 140, 100, 100, and 100 MW and the real-time offers 15, 31, 30, 10, and 20 $/MW. Lines (2, 3), (4, 5), and (6, 11) had line capacities 50, 50, and 20 MW, and other lines had no line limit. In the IEEE-118 simulation, the generators had generation costs arbitrarily selected from $\{20, 25, 30, 35, 40$ \$/MW$\}$ and generation capacities arbitrarily selected from $\{200, 250, 300, 350, 400$ MW$\}$. Total 16 lines had the line capacities arbitrarily selected from $\{70, 90, 110$ MW$\}$, and other lines had no line limit. To handle possible occurrence of price spikes, we set the upper and lower price caps as 500\$/MW and -100\$/MW respectively. Total 1000 Monte Carlo runs were executed for each case.

[††]Topology attacks need to make few additional modifications on breaker state data such that the target lines appear to be disconnected to the topology processor. However, for simplicity, we do not take into account this additional effort.

Fig. 2.3 is the plot of ARPPs[‡‡] versus detection probabilities of bad data. They show that even when bad data were detected with low probability, ARPPs were large, especially for the fully adaptive bad meter data and the bad topology data.

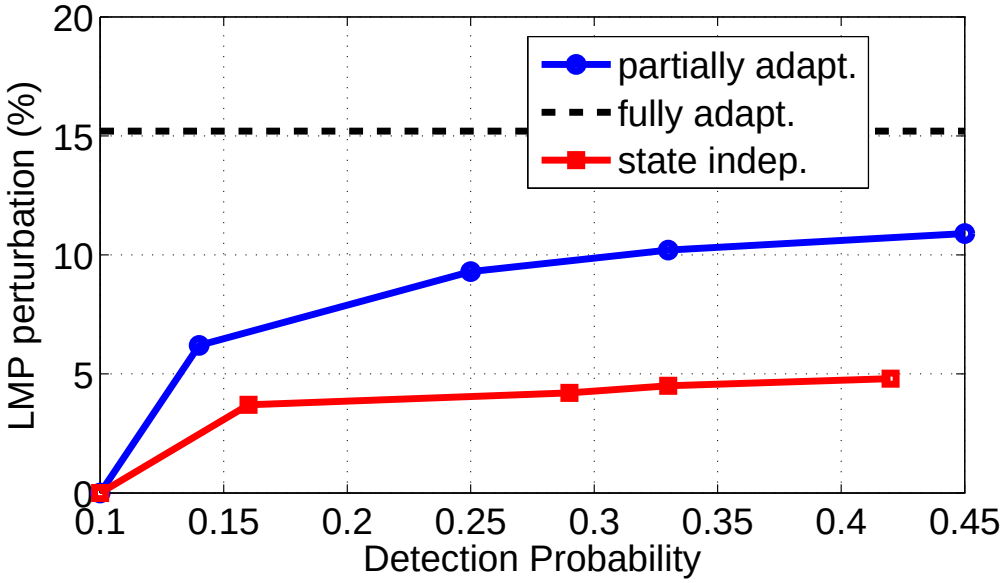
Comparing ARPPs of the three bad meter data models, we observe that the adversary may significantly improve the perturbation amount by exploiting partial or all real-time meter data (for the partially adaptive case, the adversary observed a half of all meters.) It is worthy to point out that bad topology data result in much greater price perturbation than bad meter data.

Recall the discussion in Section 2.1 and Section 2.4 that bad topology data and bad meter data employ different price-perturbing mechanisms: bad topology data perturb real-time LMP by restructuring the price regions without perturbing the state estimate (the line-removal attack introduced in Section 2.4 does not perturb state estimate) whereas bad meter data perturb real-time LMP by simply moving the state estimate to a different price region. Therefore, the observation implies that restructuring the price regions has much greater impact on real-time LMP than merely perturbing the state estimate.

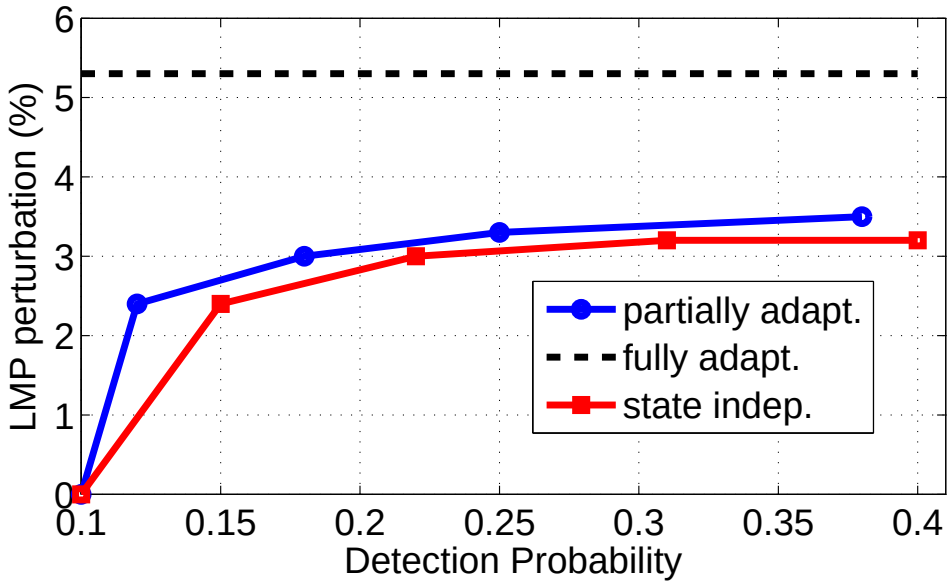
2.5.2 Nonlinear model with AC state estimation

The simulations with the nonlinear model intend to investigate the vulnerability of the real-world power system to the worst adversarial act, designed based on the linear model. The simulations were conducted on IEEE-14 and IEEE-118

[‡‡]The detection probabilities for the fully adaptive bad meter data and the bad topology data cases were less than 0.1 in all the simulations. In the figures, we draw ARPPs of those cases as horizontal lines so that we can compare them with other cases.



(a) IEEE-14: ARPP of the worst topology data is 66.1%.



(b) IEEE-118: ARPP of the worst topology data is 22.4%.

Figure 2.3: Linear model: ARPP vs detection prob.

systems in the same manner as the linear case except that we employed the nonlinear model and the AC state estimation.

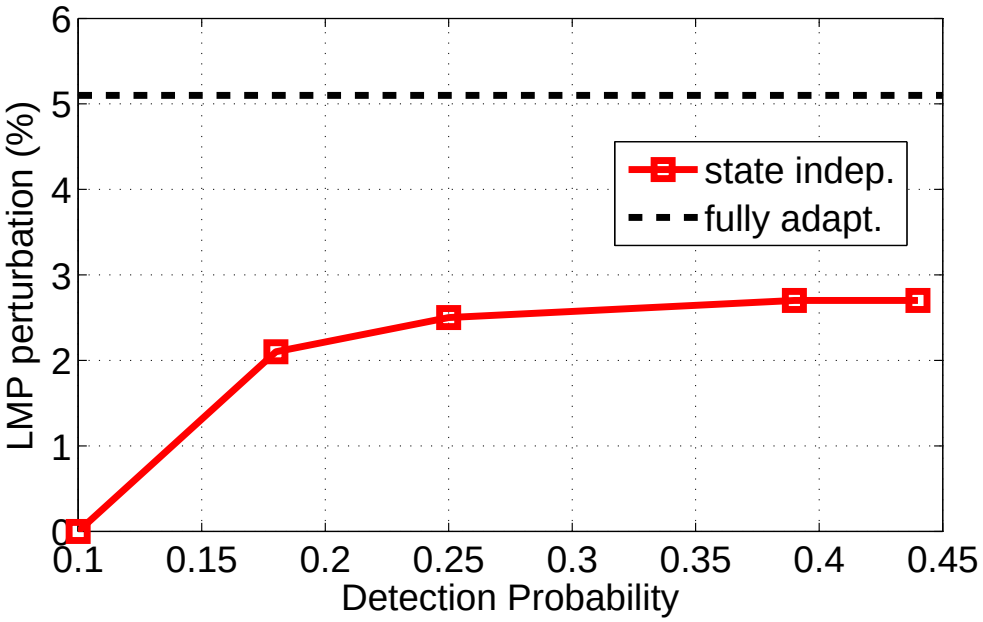
Fig. 2.4 is the plot of ARPPs versus detection probabilities. The result shows

that the proposed methodology can affect the system to some extent even when nonlinear estimator is used, especially when the bad data are present in the topology data, although the nonlinear estimator makes this effect relatively less significant compared with the linear case results.

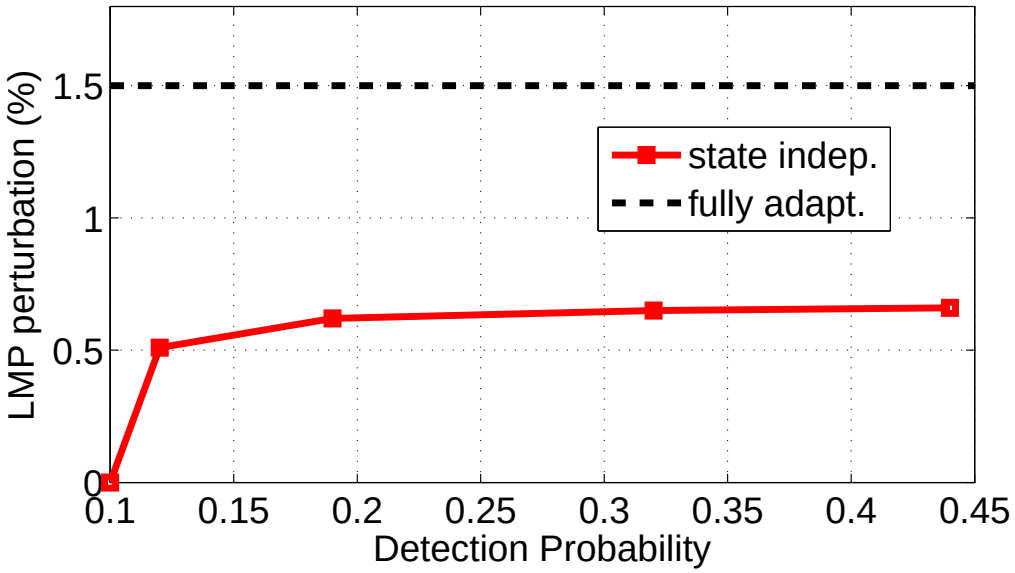
Although using nonlinear state estimation seems "unfair" to the proposed strategies, since they are all designed based on DC model, we want to emphasize two aspects here by using the simulation results. On one hand, the effects of these DC based strategies are mitigated due to the high complexity of the nonlinear model and hence the detailed structure of the nonlinear system needs to be explored; on the other hand, even though the actual nonlinear model is used, the topology attack can still achieve a very good performance.

2.5.3 Performance of the greedy search heuristic

We also conducted simulation based on the proposed greedy search technique in Section 2.3.5. The simulation was based on 118 bus system, and all parameters were the same as those presented in Section 2.5.1. We compared the performance and computation time of the greedy heuristics with exhaustive search benchmark, as shown in Table 2.1. Notice here the exhaustive search and greedy search are both over the lines that are close to their flow limits (estimated flows under M1 and M2, or actual flows under M3), the same as in Section 2.5.1. In Table 2.1, the second column (average search time) is the average searching time for worst congestion pattern over 1000 Monte Carlo runs, and the third column (accuracy) is the percentage that the greedy search find the same worst congestion pattern as the exhaustive search. From the result, we can see that using



(a) IEEE-14: ARPP of the worst topology data is 95.4%.



(b) IEEE-118: ARPP of the worst topology data is 76.9%.

Figure 2.4: Nonlinear model: ARPP vs detection prob.

greedy heuristic can give us much faster processing algorithm without losing much of the accuracy.

Table 2.1: Performance of greedy search method

method	average search time	accuracy
exhaustive search	1.23s	100%
greedy search	0.51s	97.3%

CHAPTER 3

INTERACTIONS BETWEEN ELECTRICITY RETAILER AND CONSUMERS

3.1 Stackelberg game model

We formulate the DAHP scheme as a Stackelberg game model [22]. Particularly, the retailer is the leader who makes decision first while the consumers are the followers who make decision after observing the leader's action. The action sequence in DAHP is as follows:

1. The retailer offers the consumers the day-ahead hourly retail price, taking into account the statistics of real-time wholesale market and consumer response.
2. In real-time, a consumer dynamically determines her energy consumption. The payment from a consumer to the retailer is settled as the product of the day-ahead hourly price and the real-time energy consumption.
3. The retailer meets aggregated demand by purchasing electricity at the wholesale market, possibly complemented by its own (renewable) generation resources.

The Stackelberg game can be solved via backward induction. We thus present first the analysis of optimal demand response to a fixed day-ahead hourly price vector, π. Then, the problem of optimizing π from the retailer's point of view is considered.

3.1.1 Consumer action: optimal demand response

Under the DAHP scheme, after the day-ahead hourly price vector π is given, a consumer can fully determine her own energy consumption in real time to optimize the consumer surplus (*i.e.*, the difference between consumer utility and retail electricity payment). With the help of home energy management devices, the consumer can adopt optimal control policy and the corresponding demand response can be viewed as the predicted behavior of a rational consumer in the Stackelberg game described above.

According to [2], the U.S. Energy Information Administration reports that most of the residential electricity usage comes from space heating and air conditioning. Therefore, we assume that the price elastic demand comes primarily from the control of a certain HVAC unit that maintains the indoor temperature at a certain desirable setting and we also specialize a particular thermal dynamic model involving an HVAC temperature control.

Let $\pi = (\pi_1, \cdots, \pi_{24})$ be the DAHP vector where π_i is the day-ahead price for the ith hour and $x = (x_1, ..., x_{24})$ be the vector of average indoor temperatures for each hour. Empirical study [9] has shown that the dynamic equation that governs the temperature evolution is given by

$$x_i = x_{i-1} + \alpha(a_i - x_{i-1}) - \beta p_i + w_i, \tag{3.1}$$

where $a = (a_1, a_2, ..., a_{24})$ is the vector of average outdoor temperatures in each hour, $p = (p_1, ..., p_{24})$ the vector of control variable representing the total amount of electricity drawn by the HVAC unit during each hour and $w = (w_1, w_2, ..., w_{24})$ the process noise with zero mean. System parameters α $(0 < \alpha < 1)$ and β model the insolation of the building and efficiency of the HVAC unit. Note

that the above equation applies to both heating and cooling scenarios but not simultaneously. We focus herein the cooling scenario ($\beta > 0$) and the results apply to heating ($\beta < 0$) as well.

To observe the state x and control the HVAC unit, temperature measurements need to be collected. We assume that thermometers are installed for both indoor and outdoor temperatures. The measurement equation is given by

$$y_i = (x_i, a_i) + v_i, \tag{3.2}$$

where v_i is the vector of measurement noise.

For a particular consumer j, assume that at hour i, she wants to keep the indoor temperature close to her own desired temperature $t_i^{(j)}$. In the following, we will use superscript (j) to denote the variable associated with consumer j. The deviation of the actual indoor temperature $x_i^{(j)}$ from desired temperature $t_i^{(j)}$ can be used to measure the consumer j's uncomfort level. By assuming symmetric upward and downward discomfort, a quadratic form of consumer utility as a function of indoor temperature $x^{(j)}$ is given by

$$u^{(j)} = -\mu^{(j)} \sum_{i=1}^{24} (x_i^{(j)} - t_i^{(j)})^2, \tag{3.3}$$

where $\mu^{(j)}$ is a consumer j's own weight factor to convert the squared temperature deviation to a monetary value. Note that the state variable $x^{(j)}$ is affected by the HVAC energy consumption vector $p^{(j)}$.

On the other hand, given the retail price π and the consumers' responsive demand $p^{(j)}$, the real-time payment of energy consumption from the consumer to the retailer is settled as $\sum_{i=1}^{24} \pi_i p_i^{(j)}$. Therefore, the consumer surplus, $\mathsf{cs}^{(j)}(\pi, \omega)$, can be defined as the difference between the consumer utility and total payment

from the consumer j to the retailer, as shown below. Hereafter, we use ω to indicate random variables and vectors.

$$\mathsf{cs}^{(j)}(\pi, \omega) \stackrel{\Delta}{=} u^{(j)} - \pi^{\mathrm{T}} p^{(j)}. \tag{3.4}$$

As our assumption, a rational consumer will dynamically manage her energy consumption to maximize the consumer surplus in real time according to π and current state. Therefore, the solution to the following stochastic control program serves as the optimal residential demand response to DAHP.

$$\begin{aligned} \max_{p^{(j)}} \quad & \mathbb{E}\left\{\sum_{i=1}^{24}[-\mu^{(j)}(x_i^{(j)} - t_i)^2] - \pi^{\mathrm{T}} p^{(j)}\right\} \\ \text{s.t.} \quad & x_i^{(j)} = x_{i-1}^{(j)} + \alpha^{(j)}(a_i - x_{i-1}^{(j)}) - \beta^{(j)} p_i^{(j)} + w_i^{(j)} \\ & y_i^{(j)} = (x_i^{(j)}, a_i) + v_i^{(j)}. \end{aligned} \tag{3.5}$$

Under mild conditions where the price doesn't vary too much during a day and μ is large, we ignore the positive constraint and rate constraint for energy consumption $p^{(j)}$ for computation convenience. With this simplification, the optimization problem (3.5) is similar to the standard linear quadratic Gaussian (LQG) problem where the costs on the control and states are both quadratic. Soving the problem by backward induction will give the optimal demand response as shown in the following theorem [29].

Theorem 2 *Under DAHP scheme and the assumption that solutions to (3.5) are all positive for all consumers, the optimal aggregated residential demand response for a fixed retail price π has the following affine form,*

$$d(\pi, \omega) = -G\pi + b(\omega), \tag{3.6}$$

where G is a deterministic and positive definite matrix, depending only on the dynamic system parameters.

Proof 2 *For consumer j, the optimal control is given as,*

$$\begin{aligned} p_i^{*(j)} &= \tfrac{1}{\beta}\left(\hat{x}_{i-1|i-1}^{(j)} + \alpha^{(j)}(\hat{a}_{i|i-1} - \hat{x}_{i-1|i-1}^{(j)}) - x_i^{*(j)}\right), \\ x_i^{*(j)} &\triangleq \tfrac{\pi_i - (1-\alpha^{(j)})\pi_{i-1}}{2\mu^{(j)}\beta^{(j)}} + t_i^{(j)}, \end{aligned}$$

where $\hat{x}_{i-1|i-1}^{(j)}$ and $\hat{a}_{i|i-1}$ are the ML estimated indoor and outdoor temperatures of hour i at hour $i-1$, π_{25} is assumed to be 0, and $x_i^{*(j)}$ is an ancillary value. Expanding the solution above, we will have the total demand of consumer j is

$$d^{(j)} = p^{(j)} + \tilde{p}^{(j)} = -G^{(j)}\pi + c^{(j)} + \tilde{p}^{(j)},$$

where $c^{(j)}$ is a random variable and independent of π, $\tilde{p}^{(j)}$ is the price inelastic demand from consumer j and $G^{(j)}$ satisfies

$$G_{ik}^{(j)} = \begin{cases} [1 + (1-\alpha^{(j)})^2]/[2\mu(\beta^{(j)})^2] & \text{if } i = k \neq 1 \\ 1/[2\mu(\beta^{(j)})^2] & \text{if } i = k = 1 \\ (-1 + \alpha^{(j)})/[2\mu(\beta^{(j)})^2] & \text{if } |i-k| = 1 \\ 0 & \text{o.w.} \end{cases}$$

Notice that $G^{(j)}$ is deterministic and diagonal dominant with positive diagonal elements. Hence, $G^{(k)}$ is positive definite. On the other hand, the aggregated demand

$$p(\pi, \omega) = \sum_k d^{(j)} = -G\pi + b(\omega),$$

where $b(\omega) = \sum_j (c^{(j)} + \tilde{p}^{(j)})$, $G = \sum_j G^{(j)}$. G is positive definite and deterministic, depending only on the dynamic system parameters.

Theorem 2 establishes an affine relationship between the optimal demand response and day-ahead hour price under some mild conditions. It is shown that the sensitivity matrix of demand with respect to price, $-G$, is not affected by the realization of randomness. This means that the change of price will have deterministic effect on the expected value of demand.

The result will guide our following calculation of the retailer's decision. The properties that the relationship is affine and G is positive definite are important to our later discussion.

3.1.2 Retailer action: optimal dynamic pricing

In the following, we assume that the retailer is a price taker in the wholesale market. This means that aggregated demand from the consumers would not affect the wholesale price in real time. Additionally, we assume that the Stackelberg game discussed in this chapter is with perfect information, which means that the form of the followers' (consumers') payoff function is completely known to the leader (retailer).

In real-time, the retailer is required to keep the balance of the power flow and provide aggregated demand to the consumers. No curtailment is permitted. To deliver energy, the retailer has to pay for the retail cost, including the distribution loss, the real-time payment to the wholesale market, and etc..

Let $\lambda(\omega) = (\lambda_1(\omega), \lambda_2(\omega), ..., \lambda_{24}(\omega))$ denote the random vector of average per unit retail cost during each hour. We assume there is no correlation between the demand response and the retail cost in the following. Therefore, without self-owned generation, the total daily retail profit $\mathsf{rp}(\pi, \omega)$, defined as the difference between the real-time retail revenue and the retail cost, is as follows,

$$\mathsf{rp}(\pi, \omega) = \pi^{\mathrm{T}} d(\pi, \omega) - (\lambda(\omega))^{\mathrm{T}} d(\pi, \omega). \tag{3.7}$$

The retailer's pricing decision depends on its own payoff function. If the retailer only focuses on the expected retail profit, as in a monopoly retail structure,

the solution to the following problem is the optimal pricing strategy.

$$\max_{\pi} \overline{\mathsf{rp}}(\pi) = (\pi - \bar{\lambda})^{\mathrm{T}}(-G\pi + \bar{b}), \tag{3.8}$$

where bar is used to represent the expected value, also in the following part of this chapter. As shown in Theorem 2, G is positive definite. Therefore, the problem is a convex program and can be easily solved.

However, as a load serving entity, the retailer needs to also take into consideration its consumers' satisfaction measured by the consumer surplus. Given the retail price π, by replacing the optimal demand response in Theorem 2 back into the consumer optimization problem, the expected aggregated consumer surplus can be expressed as

$$\begin{aligned} \overline{\mathsf{cs}}(\pi) &= \textstyle\sum_j \overline{\mathsf{cs}}^{(j)}(\pi) \\ &= \textstyle\sum_j \sum_{i=1}^{24} [-\mu(x_i^{*(j)} - t_i)^2] - (\bar{d}^{(j)})^{\mathrm{T}}\pi \\ &= \pi^{\mathrm{T}} G\pi/2 - \pi^{\mathrm{T}}\bar{b} + c, \end{aligned} \tag{3.9}$$

where $x^{*(j)}$ and $d^{(j)}$ are the same as the values in the proof of Theorem 2 in the Appendix, and c is a constant depending on the variance of the noise.

The other extreme payoff function the retailer may take is the expected social welfare. Formally, the expected social welfare, $\overline{\mathsf{sw}}(\pi)$, can be defined as the sum of consumer surplus and retail profit, *i.e.*,

$$\overline{\mathsf{sw}}(\pi) = \overline{\mathsf{rp}}(\pi) + \overline{\mathsf{cs}}(\pi). \tag{3.10}$$

The social welfare reflects the combined benefit of the consumers and the retailer. Maximizing the expected social welfare, we can get the following theorem.

Theorem 3 *The optimal retail price π^{sw} that maximizes the social welfare is the expected real-time retail cost, i.e.,*

$$\pi^{sw} = \bar{\lambda},$$

and the expected retail profit under π^{sw} is

$$\overline{rp}(\pi^{sw}) = 0.$$

For any π' such that $\overline{rp}(\pi') \geq 0$, we have $\overline{cs}(\pi') \leq \overline{cs}(\pi^{sw})$.

Proof 3 *Setting the derivative of $\overline{sw}(\pi)$ to zero gives the optimal price and resulted retail profit as,*

$$\pi^{sw} = \bar{\lambda}, \ \ \overline{rp}(\pi^{sw}) = 0.$$

For any π' such that $\overline{rp}(\pi') \geq 0$, we have

$$\overline{cs}(\pi') = \overline{sw}(\pi') - \overline{rp}(\pi') \leq \overline{sw}(\pi^{sw}) - 0 = \overline{cs}(\pi^{sw}).$$

Theorem 3 shows that, if the social welfare is to be maximized, the retailer generates no profit. This result is consistent with the situation when there is perfect competition among identical retailers, in which case, social welfare maximization leads to zero profit. It is also shown that when social welfare is the payoff function, the retailer simply matches the DAHP with the expected real-time retail cost.

3.2 Analysis of equilibria: achievable tradeoff

In this section, we will solve the Stackelberg game with a general form of the retailer's payoff function and establish the trade-off between consumer surplus (CS) and retail profit (RP).

As a load serving entity, the utility needs to take into consideration its own profit and consumers' satisfaction measured by the consumer surplus. In particular, we consider a weighted social welfare in expectation defined as a weighted sum of consumer surplus and retail profit, *i.e.*,

$$\max \{\overline{\mathsf{rp}}(\pi) + \eta\overline{\mathsf{cs}}(\pi)\}, \tag{3.11}$$

where η is the retailer's preference weight on the consumer surplus. If $\eta = 1$, this is equivalent to optimizing the social welfare. If $\eta = 0$, this is the same as optimizing the retail profit. When $\eta < 0$, the retailer can benefit from reducing the consumer's surplus. On the other side, when $\eta > 1$, the retailer will reduce its own profit (maybe to negative) to achieve better payoff comparing with the social welfare maximization case. Therefore, a rational retailer should choose η between $[0, 1]$ and we will not consider those η's beyond $[0, 1]$ which don't make economic sense for the retailer.

An alternative formulation is based on the optimal CS-RP trade-offs. In particular, we are interested in characterizing the Pareto front involving $(\overline{\mathsf{cs}}(\pi), \overline{\mathsf{rp}}(\pi))$. A point on the Pareto front can be obtained by considering a practical situation where the retailer optimizes its profit with the constraint that the consumer surplus no less than a certain level. In particular, the problem is formulated as

$$\begin{aligned} \max \quad & \overline{\mathsf{rp}}(\pi) \\ \text{s.t.} \quad & \overline{\mathsf{cs}}(\pi) \geq \tau. \end{aligned} \tag{3.12}$$

From the above optimization, the Pareto front can be traced by varying the consumer surplus level. Comparing the formulation (3.11) and 3.12), we have the following theorem.

Theorem 4 *For any specific η, if the solution in (3.11) is π^*, π^* is also a solution to (3.12) with $\tau = \overline{cs}(\pi^*)$. Varying τ in optimization (3.12) and varying η in optimization (3.11) will give the same trade-off curve between expected retail profit and expected consumer surplus.*

Proof 4 *With a particular η, assume π^* is a solution to (3.11). Let $\tau = \overline{cs}(\pi^*)$ in (3.12). Then π^* will be in the feasible set of (3.11). If there exists π', such that $\overline{rp}(\pi') > \overline{rp}(\pi^*)$, and $\overline{cs}(\pi') \geq \tau$,*

$$\overline{rp}(\pi') + \eta\overline{cs}(\pi') > \overline{rp}(\pi^*) + \eta\tau = \overline{rp}(\pi^*) + \eta\overline{cs}(\pi^*).$$

Hence, π^* is not the solution to (3.11) since π' achieves better objective value. It contradicts with the assumption. Therefore, π^* is also a solution to (3.12).

Theorem 4 implies that the two optimization problems are equivalent and give the same Pareto front. Each point on the Pareto front is attainable and corresponds to a equilibrium point in the Stackelberg game with particular payoff function. The shape of the trade-off region is characterized by the following theorem.

Theorem 5 *The Pareto front of $(\overline{cs}, \overline{rp})$ is concave and decreasing. The area above the Pareto front is infeasible for the retailer to achieve under DAHP.*

Proof 5 *For $\eta \in [0, 1]$, the solution to (3.11) is given by*

$$\pi^* = \frac{1}{2-\eta} G^{-1}[(1-\eta)\bar{b} + G\bar{\lambda}].$$

Define the resulted retail profit and consumer surplus as

$$rp^*(\eta) \stackrel{\Delta}{=} \overline{rp}(\pi^*(\eta)), cs^*(\eta) \stackrel{\Delta}{=} \overline{cs}(\pi^*(\eta)).$$

Numerical calculation shows that,

$$\frac{\partial rp^*(\eta)}{\partial cs^*(\eta)} = \frac{\frac{\partial rp^*(\eta)}{\partial \eta}}{\frac{\partial cs^*(\eta)}{\partial \eta}} = -\eta.$$

$cs^*(\eta)$ *is an increasing function of* η*. Therefore,* $\frac{\partial rp^*(\eta)}{\partial cs^*(\eta)}$ *decreases as* $cs^*(\eta)$ *increases. The curve is concave. According to Theorem 4,* $rp^*(\eta)$ *is the optimal value of (3.12) when consumer surplus is at least* $cs^*(\eta)$*. Therefore, no CS-RP pair can be above the trade-off curve.*

Some well known pricing strategies can be placed on the Pareto front, as shown in Fig. 1.2. The social welfare maximizing pricing π^{sw} is located on the CS axis. This is from Theorem 3 and also intuitive since maximizing social welfare dictates the removal of retail profit. The optimal regulated monopoly price π^r is located at the Pareto front where the retailer profit has a regulated profit margin, Δ. π^o is the price when the retailer's objective is purely maximizing its profit, which is the leftmost point on the trade-off curve.

3.3 Effect of renewable energy

With the integration of renewable energy, the characteristics of the interaction between the retailer and the consumers will be changed, *i.e.*, the CS-RP trade-off curve will be shifted. However, the effect depends on whether the renewable energy is utility-based or consumer-based.

3.3.1 Effect of utility-based wind integration

We consider in this subsection the role of renewable energy owned by the retailer. As a large load aggregator, the retailer may have the financial ability and incentive to have its own wind farm or have contract with other wind farms. We now consider a scenario that the retailer has access to wind power which can be used to compensate the real-time loads from the consumers. We assume that the usage of the wind farm is completely for serving the local area, and excess wind power will be spilled by the retailer.

In practice, the marginal cost of wind power is almost negligible. For simplicity, in the following we assume zero cost of using wind power in real time and use a random vector, $q(\omega) = (q_1(\omega), ..., q_{24}(\omega))$, to denote the wind power available for the retailer in each hour. Therefore, the retailer's profit after wind power integration is in the form of

$$\mathsf{rp}_w(\pi, \omega) = \pi^{\mathrm{T}} d(\pi, \omega) - \lambda(\omega)^{\mathrm{T}} (d(\pi, \omega) - q(\omega))^+, \tag{3.13}$$

where function $(x)^+$ is the positive part of x, defined as $(x)^+ \triangleq \max\{x, 0\}$.

Following a similar discussion as in the Section 3.2, we focus on the problem that the retailer's payoff function is a linear combination of retail profit and consumer surplus. Notice here the consumer surplus does not change while the retail profit is replaced by $\mathsf{rp}_w(\pi, \omega)$, where the preference weight on consumer surplus is denoted as $\eta \in [0, 1]$. Varying η and maximizing the weighted social welfare $\overline{\mathsf{rp}}_w(\pi) + \eta\overline{\mathsf{cs}}(\pi)$ will give the CS-RP trade-off curve with utility-based wind power.

By incorporating free wind power, intuitively the CS-RP trade-off curve will be enlarged towards upright on the CS-RP plane. The following theorem ver-

ifies this intuition and also shows how the benefit of wind integration is distributed between the retailer and the consumers.

Theorem 6 *Assume that for each hour i, the available wind power is uniformly distributed over $[0, K]$, where K denotes the maximum capacity. For a particular preference parameter η, let the optimal demand level before wind integration in the optimization (3.11) be $d(\eta)$ and the optimal price after wind integration be π_w^η. Define $\Delta\overline{rp}(\eta)$ and $\Delta\overline{cs}(\eta)$ as the increase of retail profit and consumer surplus after wind integration. Then we have,*

1. *The retail profit change is always positive, i.e., $\Delta\overline{rp}_w(\pi_w^\eta) > 0$ for all η.*
2. *When $K \leq \min_{i\in\{1,\dots,24\}} d_i(\eta)$, $\Delta\overline{cs}(\eta) = 0$. Otherwise, $\Delta\overline{cs}(\eta) > 0$.*
3. *As K goes to ∞, the fraction of wind integration benefit to the consumer side, $\frac{\Delta\overline{cs}(\eta)}{\Delta\overline{cs}(\eta)+\Delta\overline{rp}(\eta)}$, goes to $\frac{1}{3-2\eta}$.*

Proof 6 *Before wind integration, for a particular η, the first order condition gives that the optimal demand level $d(\eta)$ satisfies*

$$b - (2-\eta)d(\eta) = G\lambda.$$

After wind integration, for a particular η, the first order condition gives that the optimal demand level $d_w(\eta)$ satisfies

$$b - (2-\eta)d_w(\eta) = G(\lambda \circ F(d_w(\eta))),$$

where $\circ$ means the Hadamard product, i.e., piecewise product of two vectors, and F is the cdf of wind power distribution.

When $K < \min_i d_i(\eta)$, we can see that $d(\eta)$ satisfies the optimal condition therefore $d_w(\eta) = d(\eta)$, $\Delta\overline{cs}(\eta) = 0$. Otherwise, $d_w(\eta) = d(\eta) + G\delta/(2-\eta)$, where $\delta = \lambda - \lambda \circ$

$F(d_w(\eta)) \geq 0$, *and* δ *is not zero since* $d(\eta)$ *does not satisfy the optimal condition after wind integration.*

$$\begin{aligned}\Delta\overline{cs}(\eta) &= \tfrac{1}{2}\{(d_w(\eta))^T G^{-1} d_w(\eta) - (d(\eta))^T G^{-1} d(\eta)\} \\ &= \tfrac{1}{2}\{2\delta^T d(\eta) + \delta^T G\delta\} > 0.\end{aligned}$$

For the RP, for all K and η*,*

$$\begin{aligned}\Delta\overline{rp}(\eta) &= (1-\eta)\{(d_w(\eta))^T G^{-1} d_w(\eta) - (d(\eta))^T G^{-1} d(\eta)\} \\ &\quad + \tfrac{1}{2K}\{(d_w(\eta))^T \Lambda d_w(\eta)\} > 0,\end{aligned}$$

where $\Lambda = diag(\bar{\lambda}_1, ...\bar{\lambda}_{24})$*. As K goes to infinity,* $d_w(\eta)$ *is bounded,* $\Delta\overline{rp}(\eta)$ *goes to* $(1-\eta)\{(d_w(\eta))^T G^{-1} d_w(\eta) - (d(\eta))^T G^{-1} d(\eta)\}$*,* $\Delta\overline{cs}(\eta)$ *equals to* $\frac{1}{2}\{(d_w(\eta))^T G^{-1} d_w(\eta) - (d(\eta))^T G^{-1} d(\eta)\}$*, then* $\frac{\Delta\overline{cs}(\eta)}{\Delta\overline{cs}(\eta)+\Delta\overline{rp}(\eta)}$ *goes to* $\frac{1}{3-2\eta}$*.*

In Theorem 6, it is shown that the retail profit change with wind power integration is always positive. In particular, for social welfare maximization, the retailer profit is positive and the corresponding point is economically viable.

Theorem 6 also shows that the benefit of wind integration all goes to the retailer side when the capacity is small. As the capacity of wind farm is larger than a certain threshold, the fraction of the wind integration benefit to the consumer side becomes positive. When the capacity goes to infinity, this fraction converges to a particular limit, $\frac{1}{3-2\eta}$, depending only on the retailer's preference weight η over the consumer surplus.

The convergence limit in Theorem 6 implies that when the retailer is purely a "profit seeker", the fraction of wind integration to demand side converges to $\frac{1}{3}$. On the other hand, if the retailer's objective is social welfare maximization, as the capacity of wind power goes to infinity, the fraction of renewable integration benefit to consumers converges to 1, which means that the retail profit will converge to zero.

3.3.2 Effect of consumer-based renewable integration

On the other hand, the consumers may also have access to renewable energy, e.g., solar energy. Nowadays, residential solar panels are widely used and solar energy can be viewed as an important substitute for the electricity purchased directly from the retailer. In the following, we will study the effects of consumer-based solar energy and denote it as a 24-dimensional random nonnegative vector, $s(\omega) = (s_1(\omega), s_2(\omega), ..., s_{24}(\omega))$, with the mean $\bar{s}$. Under DAHP structure, the consumer's optimization problem (3.5) is changed to

$$\begin{aligned} \max_{p,r} \quad & \mathbb{E}\left\{\sum_{i=1}^{N}[-\mu(x_i - t_i)^2] - \pi^{\mathrm{T}}(u - s)\right\} \\ \text{s.t.} \quad & x_i = x_{i-1} + \alpha(a_i - x_{i-1}) - \beta u_i + w_i, \\ & y_i = (x_i, a_i) + v_i \end{aligned} \tag{3.14}$$

Following the same backward induction in Theorem 2, we can see that given the same day-ahead price π, the HVAC energy usage u will not change, while the aggregated demand $d_s(\pi, \omega)$ has the form,

$$d_s(\pi, \omega) = b(\omega) - s(\omega) - G\pi. \tag{3.15}$$

Accordingly, the expected consumer surplus with solar energy is changed to $\overline{\mathsf{CS}}_s(\pi) = \overline{\mathsf{CS}}(\pi) + \pi^{\mathrm{T}}\bar{s}$, and the retail profit $\overline{\mathsf{RP}}_s(\pi) = \overline{\mathsf{RP}}(\pi) - (\pi - \bar{\lambda})^{\mathrm{T}}\bar{s}$. With different preference weight on consumer surplus, $\eta \in [0, 1]$, maximizing the retail payoff function, $\overline{\mathsf{RP}}_s(\pi) + \eta\overline{\mathsf{CS}}_s(\pi)$, will give us the CS-RP tradeoff curve with utility-based solar energy. The following theorem characterizes the shape of the new trade-off curve.

Theorem 7 *Denote the social welfare and retail profit maximization price as π_s^o and π_s^{sw}, respectively.*

1. *The social welfare maximization price is $\bar{\lambda}$, and the resulted retail profit, $\overline{rp}_s(\bar{\lambda})$ is 0. The corresponding social welfare maximization point $(\overline{cs}_s(\pi_s^{sw}), \overline{rp}_s(\pi_s^{sw}))$ is outside the original CS-RP trade-off curve without solar energy.*
2. *The maximized retail profit $\overline{rp}_s(\pi_s^o))$ is smaller than the maximized retail profit without consumer-based solar energy.*
3. *As $\bar{s}$ increases, the maximized social welfare increases and the maximized retail profit decreases.*

Proof 7 *When $\eta = 1$, the retailer's payoff function is social welfare,*

$$\overline{sw}_s(\pi) = \overline{sw}(\pi) + \bar{\lambda}^T \bar{s}. \tag{3.16}$$

Therefore, the social welfare maximization price will be the same as in Theorem 3, $\bar{\lambda}$, and the resulted retail profit, $\overline{rp}_s(\bar{\lambda})$ is still 0. Since $\bar{\lambda}^T \bar{s} > 0$, the social welfare maximization point, $(0, \overline{sw}_s(\bar{\lambda}))$, will be outside the original CS-RP trade-off curve without solar energy.

When $\eta = 1$, the retailer's payoff function is retail profit, and the optimal price is calculated as

$$\pi_s^o = \frac{1}{2}(G^{-1}\bar{b} - G^{-1}\bar{s} + \bar{\lambda}). \tag{3.17}$$

The retail profit change with solar energy is

$$\Delta\overline{rp} = -\frac{1}{2}(G^{-1}b - \lambda)^T \bar{s}, \tag{3.18}$$

Since $G^{-1}b$ is the price to make demand equal to zero, $G^{-1}b - \lambda \geq 0$. Therefore, $\Delta\overline{rp} < 0$. The maximized profit decreases.

Also, the maximized RP decreases with the increase of solar energy level $\bar{s}$.

■

Theorem 7 shows that on the left side of the CS-RP trade-off curve, the maximized profit with solar energy is less than previous maximized profit and the CS-RP pairs with solar energy are below the original trade-off curve. While on the right side, the social welfare maximization point with solar energy is outside the original trade-off curve. Therefore, two CS-RP trade-off curves, with and without solar energy, will cross with each other. Let's assume the cross point is $(\overline{cs}^*, \overline{rp}^*)$.

First we consider the case that the original operating point on the trade-off curve is to the left of $(\overline{cs}^*, \overline{rp}^*)$. After incorporating solar energy, either the retailer can no longer have the retail profit as much as before or the consumer surplus has to be reduced to maintain the same amount of retail profit. Therefore, the retail market will suffer loss even though solar energy is provided.

On the other hand, if the original operating point on the trade-off curve is to the right of $(\overline{cs}^*, \overline{rp}^*)$, the consumer surplus will increase if the retail profit remains the same. The retail market benefits from consumer-based solar energy. These results are illustrated by simulations in Section 3.5.4.

3.4 Effect of storage

With the development of energy storage technology, either the retailer and the consumers may install storage devices, which can help to reschedule the energy usage and result in better payoff. Similar to Section 3.3, these two cases also have significantly different characteristics. In the following, we will study

the effects of utility-based and consumer-based storage on the CS-RP trade-off curves separately.

3.4.1 Effect of utility-based storage

Firstly, we consider the role of storage on the utility side. For hour i, denote the energy level in the battery as B_i and the energy charged into the battery as r_i (when $r_i \leq 0$, it means discharging the battery). We use $r_i^+ \geq 0$ and $r_i^- \geq 0$ to represent the positive and negative part of r_i, *i.e.*, $r_i = r_i^+ - r_i^-$. Then the dynamics of the battery can be expressed as

$$B_{i+1} = \kappa(B_i + \tau r_i^+ - r_i^-/\rho). \tag{3.19}$$

where $\kappa \in (0,1)$ is the storage efficiency, $\tau \in (0,1)$ the charging efficiency and $\rho \in (0,1)$ the discharging efficiency.

Given the day-ahead price, π, the demand side optimization (3.5) will not change and the optimal demand is still a linear function of π as in (3.6). On the other hand, by using storage, the retailer can take advantage of hourly varying retail cost, λ, to achieve better profit. Formally, assuming the storage size is smaller than the demand, the retail profit with utility-based storage, $\overline{\mathsf{rp}}_{\mathrm{rb}}(\pi)$, can be expressed as,

$$\begin{aligned} \overline{\mathsf{rb}}_{\mathrm{rb}}(\pi) = \max_{r,B} \quad & \pi^{\mathrm{T}}\bar{d} - \bar{\lambda}^{\mathrm{T}}(\bar{d} + r) \\ \text{s.t.} \quad & \bar{d} = \bar{b} - G\pi \\ & B_{i+1} = \kappa(B_i + \tau r_i^+ - r_i^-/\rho) \\ & B_{24} = B_0, 0 \leq B_i \leq C \\ & 0 \leq r_i^+ \leq r^{\mathrm{u}}, 0 \leq r_i^- \leq r^{\mathrm{d}} \end{aligned} \tag{3.20}$$

where B_0 is the initial energy level in the storage, C the capacity of the battery, r^{u} the charging limit, and r^{d} the discharging limit.

Notice that $\overline{\mathsf{rp}}_{\mathrm{rb}}(\pi)$ can be expressed as sum of the original retail profit, $\overline{\mathsf{rp}}(\pi)$, and the arbitrage profit $Q(\bar{\lambda})$, *i.e.*, $\overline{\mathsf{rp}}_{\mathrm{rb}}(\pi) = \overline{\mathsf{rp}}(\pi) + Q(\bar{\lambda})$, where Q is defined as,

$$\begin{aligned} Q(\bar{\lambda}) \triangleq \max_{r,B} \quad & -\bar{\lambda}^{\mathrm{T}} r \\ \text{s.t.} \quad & B_{i+1} = \kappa(B_i + \tau r_i^+ - r_i^-/\rho) \\ & B_{24} = B_0, 0 \le B_i \le C \\ & 0 \le r_i^+ \le r^{\mathrm{u}}, 0 \le r_i^- \le r^{\mathrm{d}} \end{aligned} \tag{3.21}$$

With the preference weight parameter on consumer surplus η, the retailer's payoff function is

$$\overline{\mathsf{rp}}_{\mathrm{rb}}(\pi) + \eta\overline{\mathsf{cs}}(\pi) = \overline{\mathsf{rp}}(\pi) + \eta\overline{\mathsf{cs}}(\pi) + Q(\lambda). \tag{3.22}$$

Therefore, for any η, the optimal price for the retailer is the same as in the case without storage, and the CS-RP trade-off curve is parallelly shifted up by $Q(\bar{\lambda})$.

3.4.2 Effect of consumer-based storage

Then, we consider the effects of consumer-based storage. In particular, we assume the net-metering option where a consumer can, in effect, sell back excess energy. In the presence of storage, a consumer can take advantage of the hourly varying day-ahead retail price.

Using the the same notation for storage as in Section 3.4.1, the optimal de-

mand response problem on the consumer side (3.5) is changed to

$$\begin{aligned} \max_{p,r,B} \quad & \mathbb{E}\left\{\sum_{i=1}^{N}[-\mu(x_i - t_i)^2] - \pi^{\mathrm{T}}(p + r)\right\} \\ \text{s.t.} \quad & x_i = x_{i-1} + \alpha(a_i - x_{i-1}) - \beta p_i + w_i, \\ & y_i = (x_i, a_i) + v_i \\ & B_{i+1} = \kappa(B_i + \tau r_i^+ - r_i^- / \rho) \\ & B_{24} = B_0, 0 \le B_i \le C \\ & 0 \le r_i^+ \le r^{\mathrm{u}}, 0 \le r_i^- \le r^{\mathrm{d}} \end{aligned} \tag{3.23}$$

Under the net-metering assumption (see [70] for more general results), we can see that the optimization problem can be divided into two independent sub problems, where the first one is the same as the previous optimal stochastic HVAC control (3.5) and the second one is purely energy arbitrage. This means that adding storage on the demand side doesn't change the original linear relationship between the actual HVAC consumption and retail price; the benefit of storage goes the consumer side in the form of arbitrage options.

Therefore, given day-ahead price π, denote $r(\pi)$ as the optimal charging vector, *i.e.*, the solution to getting $Q(\pi)$ as in (3.21). Correspondingly, the consumer surplus with consumer-based storage is $\overline{\mathsf{cs}}_{\mathrm{cb}}(\pi) = \overline{\mathsf{cs}}(\pi) + Q(\pi)$, and the retail profit with consumer-based storage is $\overline{\mathsf{rp}}_{\mathrm{cb}}(\pi) = \overline{\mathsf{rp}}(\pi) - Q(\pi) - \overline{\lambda}^{\mathrm{T}} r(\pi)$ assuming the aggregated storage size is smaller than demand. By varying $\eta \in [0, 1]$, maximizing the retail payoff function will give us the tradeoff curve between CS and RP with consumer-based storage.

The following theorem characterizes the shape of the new CS-RP trade-off curve with consumer-based storage.

Theorem 8 *Denote the social welfare and retail profit maximization price as* π^o_{cb} *and*

π_{cb}^{sw}, *respectively.*

1. *The social welfare maximization price is* $\bar{\lambda}$*, and the resulted retail profit,* $\overline{rp}_{cb}(\bar{\lambda})$ *is 0. The corresponding social welfare maximization point* $(\overline{cs}_{cb}(\pi_{cb}^{sw}), \overline{rp}_{cb}(\pi_{cb}^{sw}))$ *is outside the original CS-RP trade-off curve without storage.*
2. *The maximized profit* $\overline{rp}_{cb}(\pi_{cb}^{sw}))$ *is less than the original maximized profit if B is small enough and* $Q(\tilde{\pi}) > 3Q(\bar{\lambda})$*, where* $\tilde{\pi} = G^{-1}b$*, which is the cut-off price resulting in zero demand.*
3. *Assuming the same storage parameters, the CS-RP trade-off curve with consumer-based storage is always inside the trade-off curve with utility-based storage.*

Proof 8 *When the preference weight factor on consumer surplus* $\eta = 1$*, the retailer's payoff function is social welfare. Notice that*

$$\begin{aligned} \overline{rp}_{cb}(\bar{\lambda}) &= \overline{rp}_{cb}(\bar{\lambda}) + \overline{cs}_{cb}(\bar{\lambda}) \\ &= \overline{rp}(\bar{\lambda}) + \overline{cs}(\bar{\lambda}) + Q(\bar{\lambda}) \\ &\geq \overline{rp}(\pi) + \overline{cs}(\pi) - \bar{\lambda}^T r(\pi), \end{aligned} \tag{3.24}$$

for any π*, which means that the social welfare maximization price is* $\bar{\lambda}$*, resulted retail profit is 0, and the maximized social welfare is increased by* $Q(\lambda)$*. Therefore, the social welfare maximization point is outside the original trade-off curve.*

When the preference weight factor $\eta = 0$*, the retailer's payoff function is retail profit. With the consumer-based storage, the consumer surplus increases by* $Q(\pi_{cb}^{o})$*, while the retail profit is changed by* $Q(\pi_{cb}^{o}) - \bar{\lambda}^T r(\pi_{cb}^{o}))$*. We know* $Q(\pi)$ *is a piecewise linear function of* π*. Therefore,*

$$
\begin{aligned}
Q(\pi^o_{cb}) &= Q((\tilde{\pi} + \bar{\lambda})/2 + \Delta\pi) \\
&\geq Q(\tilde{\pi})/2 - max_\pi(-\bar{\lambda}^T r(\pi))/2 - max_\pi(-(\Delta\pi)^T r(\pi)) \\
&\geq Q(\bar{\lambda}) + (Q(\tilde{\pi}) - 3Q(\bar{\lambda}))/2 - max_\pi(-(\Delta\pi)^T r(\pi))
\end{aligned} \tag{3.25}
$$

According to the assumption $Q(\tilde{\pi}) > 3Q(\bar{\lambda})$), when the storage size is small enough to make $\Delta\pi$ small enough, $Q(\pi^o_{cb}) > Q(\bar{\lambda})$. The retail profit decreases. We have the maximized profit decreased when the storage size is small.

For a particular η, assume the optimal price with consumer-based storage is $\pi_{cb}(\eta)$. Using the same price for the case of utility-based storage. The change of social welfare is $Q(\bar{\lambda}) - (-\bar{\lambda}^T r(\pi_{cb}(\eta)) \geq 0$. The corresponding point has decreased consumer surplus and increased retail profit. Since all the left points have less social welfare, this point is outside the trade-off curve with consumer-based storage. On the other hand, $\pi_{cb}(\eta)$ is a feasible price for the case with utility-based storage and the corresponding point is inside the trade-off curve with utility-based storage. Therefore, the utility-based storage results in a trade-off curve completely outside the one with consumer-based storage. ■

Theorem 8 shows that maximized retail profit with consumer-based storage decreases while the maximized social welfare increases, comparing with the case without storage. Therefore, the new trade-off curve is inside the original curve on the left side and outside the original curve on the right side; the two curves will cross with each other. This means that only when the preference η chosen by the retailer is larger than the one for the cross point, the consumer-based storage will bring benefit to both the retailer and the consumers.

Since demand level varies a lot across hours, especially for HVAC control, the cut-off price $\tilde{\pi}$ will lead to high arbitrage opportunity. Therefore, the condition $Q(\tilde{\pi}) > 3Q(\bar{\lambda})$ is reasonable and also consistent with our simulation setting

in Section 3.5.5.

Another conclusion from Theorem 8 is that utility-based storage results in an enlarged CS-RP trade-off curve comparing with the trade-off curve with consumer-based storage of the same size. This shows that without other consideration, the storage devices should be better controlled by the retailer under the monopoly retail market structure.

3.5 Numerical results

3.5.1 Parameter setting

In this section, we used the actual temperature record in Hartford, CT, from July 1st, 2012 to July 30th, 2012. The day-head price (used as prediction) and real-time price (used as realization) were also for the same period from ISO New England. The HVAC parameters for the simulation was set as: $\alpha = 0.5, \beta = 0.1$, $\mu = 0.5$. The desired indoor temperature was set to be $18^{\circ}C$ for all hours. The size of total consumers is 1000.

For wind power, we assumed that it is uniformly distributed, and there is no correlation across different hours.

For solar energy, we used the average solar radiation data [3] of Hartford, CT, from July 1st to July 31st. The maximum solar panel capacity was set to be 1.5MW.

For storage, we set the storage size as 10KW for indiviual houses, initial

storage level as 0, all efficiency coefficient as 0.95 and ramp limit as 5KW, and the percentage of consumers owning storage is 50%.

3.5.2 Benchmark comparisons

In order to show the efficiency of optimized DAHP and the properties of induced CS-RP tradeoff curve, we used the following two pricing schemes to compare: constant pricing and proportional mark-up pricing.

1. Constant pricing: in this case, the price remains constant for the whole day, *i.e.*, $\pi_1 = \pi_2 = ... = \pi_{24}$. By varying the price, we got the different CS-RP pairs.
2. Proportional mark-up pricing: in this case, the ratio of day-ahead price to the expected real-time price remains the same, *i.e.*, $\frac{\pi_1}{\mathbb{E}\lambda_1} = \frac{\pi_2}{\mathbb{E}\lambda_2} = ... = \frac{\pi_{24}}{\mathbb{E}\lambda_{24}}$. By varying the ratio, we got different CS-RP pairs.

The trade-off curves under these two schemes and optimal DAHP are plotted together in Fig. 3.1. Fig. 3.1 shows that the trade-off curve between CS and RP is indeed concave. When the preference weight factor η is 1, which means that the retailer's payoff function is social welfare, the resulted retail profit is zero and the corresponding point on the curve is on the CS-axis.

Comparing the three pricing schemes in Fig. 3.1, the CS-RP trade-off curves corresponding to constant and proportional mark-up pricing schemes both fall

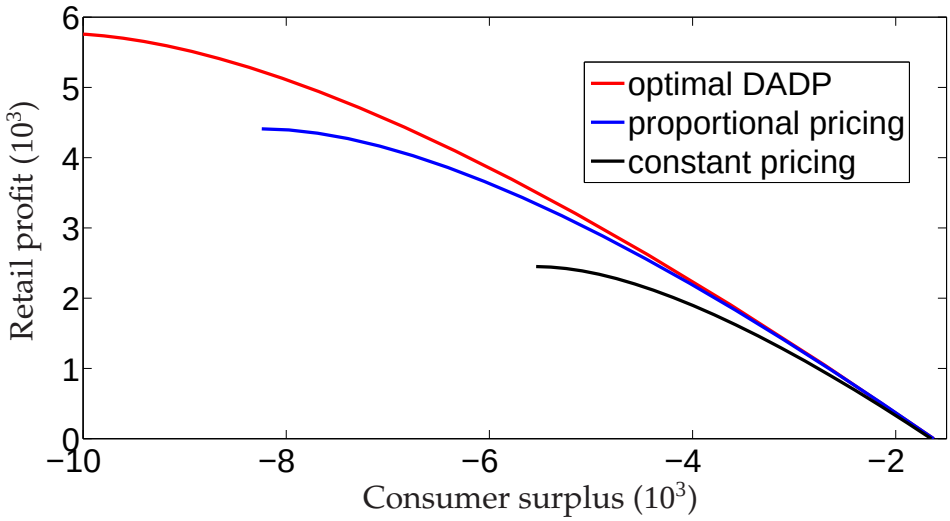


Figure 3.1: Comparasion of three pricing schemes

below the Pareto front as we expected, which shows the efficiency of the optimized DAHP.

3.5.3 Effect of utility-based wind power

Assuming the wind power in each hour is uniformly distributed over $[0, K]$, by varying preference weight on consumer surplus, η, we plotted the CS-RP tradeoff curve as shown in Fig. 3.2. The tradeoff curve is enlarged after wind power integration.

From the result shown in Fig. 3.2, we can see that when retailer integrates wind power, the social welfare optimal pricing becomes economically viable, *i.e.*, the social welfare maximization prices (rightmost points on the trade-off curve) result in positive retail profit. Furthermore, when the capacity of wind power is small (K = 1MW), the trade-off curve goes directly up rather than right, which means that almost all the benefit from wind integration goes to the

retailer side. On the other hand, when the capacity is larger (K = 3MW), the trade-off curve goes upright, which means some part of the wind integration benefit goes to consumer surplus. These results verify the statements in Theorem 6.

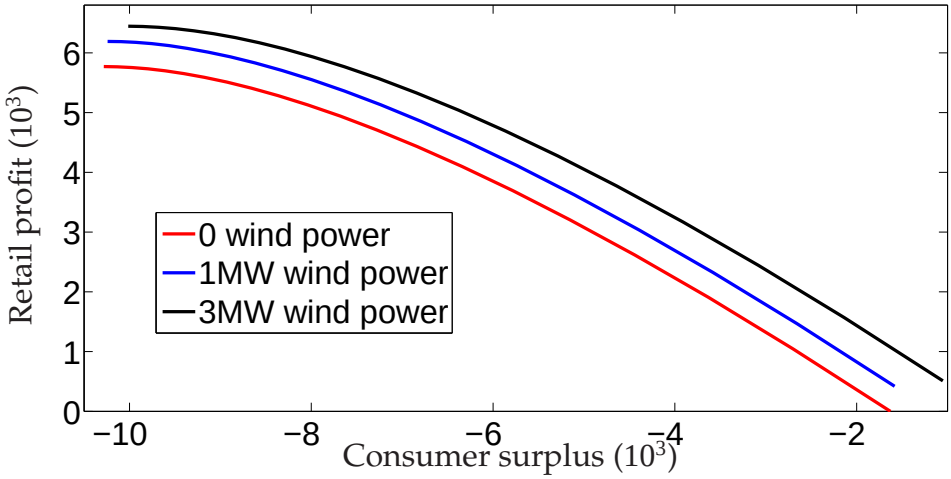


Figure 3.2: Trade-off curve with utility-based wind integration

We also plotted the distribution of wind integration benefit to the consumer side and to the retailer side with different wind integration levels, as shown in Fig. 3.3. From the result, we can see that as the level of integration (K) increases, the fraction of the wind integration benefit to CS, $\frac{\Delta cs(\eta)}{\Delta rp(\eta)}$, also increases, and converges to $\frac{1}{3-2\eta}$ as K goes to infinity, as proved in Theorem 6.

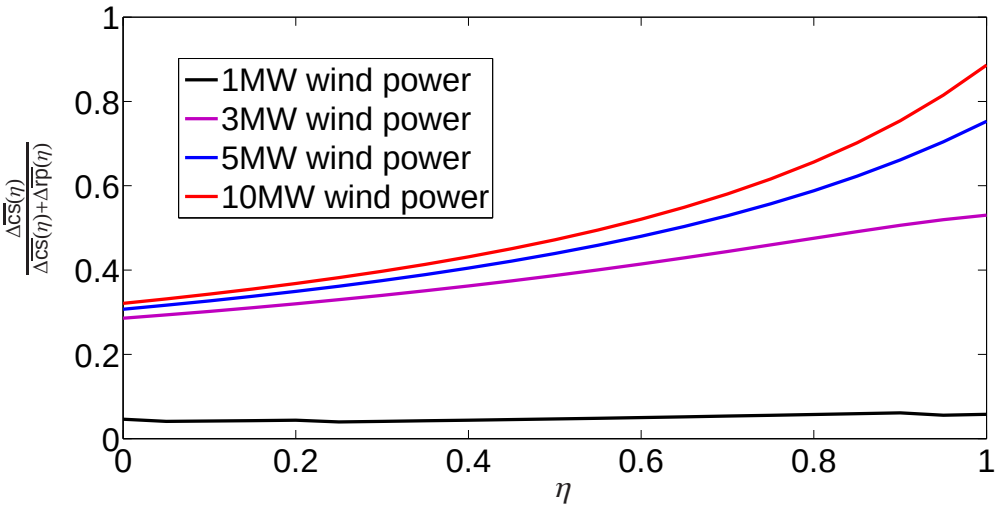


Figure 3.3: Fraction of wind benefit to consumer

3.5.4 Effect of consumer-based solar energy

Now we consider the case that consumer-based solar energy is used. Two scenarios were simulated: the percentage of consumers implement solar energy is 20% and 50%. We plotted these two CS-RP trade-off curves together with the original CS-RP trade-off curve as in Fig. 3.4.

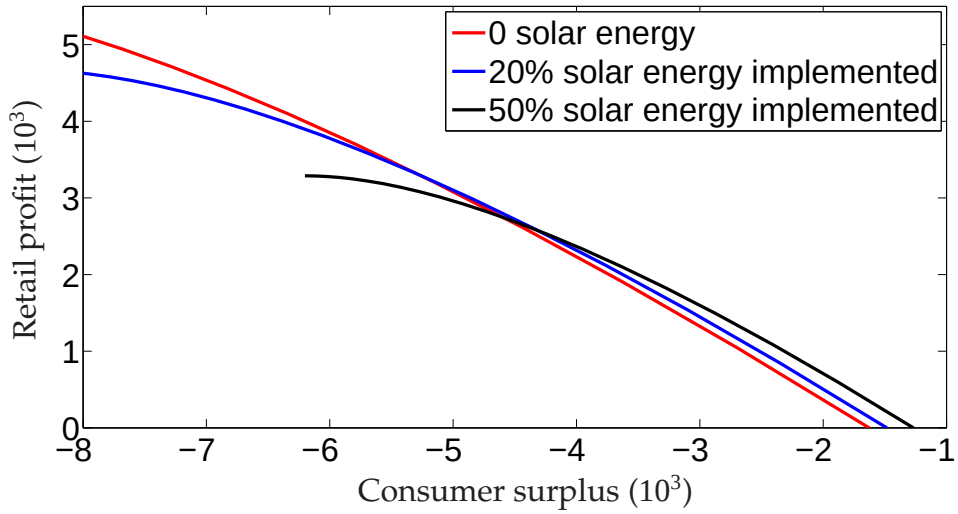


Figure 3.4: CS-RP trade-off curve with consumer-based solar energy

Fig. 3.4 shows that with solar energy, the new CS-RP trade-off curve will cross with the original one. This means that only when the retailer is operating on the right of the cross point in the beginning, having solar energy can add benefit to both sides of the retail market.

On the other hand, as the solar energy size increases, the maximized retail profit decreases, and the cross point moves to the right along the original trade-off curve. Therefore, the threshold η, which makes adding solar energy beneficial to both sides, increases. However, the maximized social welfare increases as the size of solar energy increases.

Furthermore, we compared distributed and centralized control of solar energy with the same level of integration. As shown in Fig. 3.5, the tradeoff curve with utility-based renewable is completely outside the one with consumer-based integration, which means that utility-based renewable brings more benefit to the retail market, if no other considerations are taken.

3.5.5 Effect of storage

As discussed in Section 3.4, we considered the effects of both utility-based and consumer-based storage devices on the CS-RP trade-off curves. With the assumption that utility-based storage has the same size as the consumer-based storage, (20% of the consumers have the storage devices), the results were plotted as in Fig. 3.6.

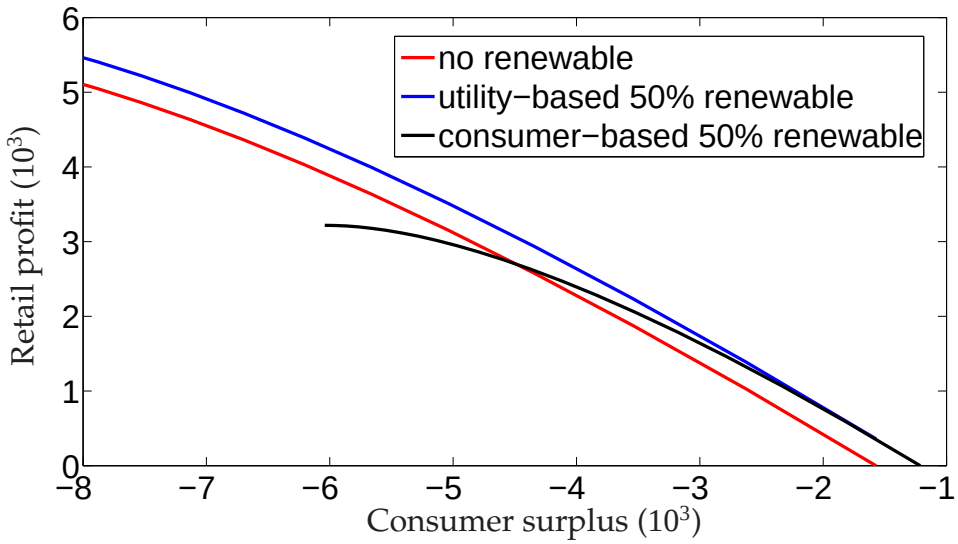


Figure 3.5: CS-RP trade-off curve comparison with utility-based and consumer-based solar energy

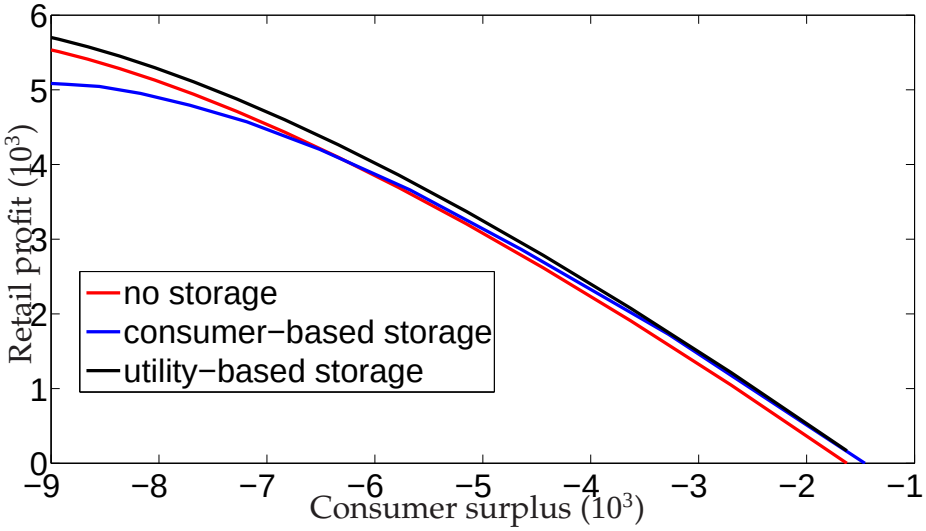


Figure 3.6: CS-RP trade-off curve with storage devices

Fig. 3.6 shows that when the retailer has access to the storage devices, the CS-RP trade-off curve is shifted upward by $Q(\bar{\lambda})$ and the social welfare point becomes economically viable.

On the other hand, as for the consumer-based storage, when the weight on consumer surplus $\eta = 1$, the retail profit remains zero, and the consumer surplus increases by $Q(\bar{\lambda})$. When $\eta = 0$, the retailer maximizes its own profit. The leftmost point on the new CS-RP trade-off curve shows that the consumer surplus increases while the retail profit decreases at a much faster rate comparing with the original tradeoff curve. In this case, the profit maximization point is inside the original CS-RP trade-off curve. According to the simulation setting, the arbitrage profit with cutoff price is much higher than with wholesale price, $Q(\tilde{\pi}) > Q(\bar{\lambda})$, *i.e.*, the condition of Theorem 8 is satisfied.

The CS-RP trade-off curve with consumer-based storage crosses with the original CS-RP trade-off curve. This means that only when the retailer is operating on the right of the cross point, consumer-based storage can benefit the retail market. On the contrary, the CS-RP trade-off curves with utility-based storage is outside the original CS-RP trade-off curve. Therefore, utility-based storage always benefits the retail market.

Comparing the trade-off curves with utility-based storage and consumer-based storage, we can see that the former one is completely outside the latter, which means that having utility-based storage brings more benefit to the retail market than consumer-based one, if no other considerations are taken.

For utility-based storage, as shown in Section 3.4, the optimal DAHP prices won't change after storage is used. As for consumer-based storage, we plotted the profit maximization prices with different storage implementation levels as shown in Fig. 3.7. Due to the arbitrage opportunity with consumer-based storage, the price becomes flatter. This shows that only consumer-based storage will change the consumers' energy consumption pattern.

Notice that almost all the end of day prices are the same for these three scenarios. The reason is that during a day, both the retail cost and temperature increase first then decrease. This means that there is little arbitrage opportunity during the second half of the day. The end of day price is affected by a very small amount.

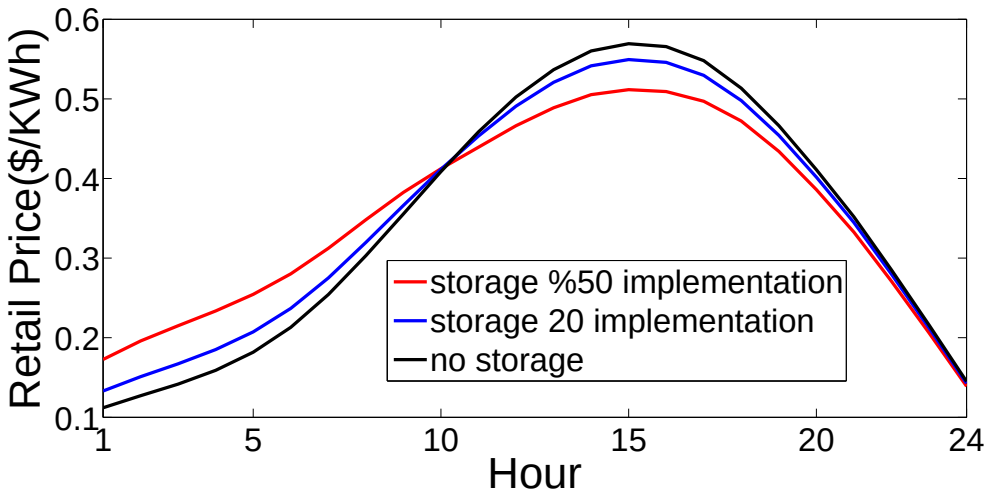


Figure 3.7: Optimal DAHP with storage devices

CHAPTER 4

ONLINE LEARNING OF DYNAMIC RETAIL PRICE

4.1 Structure of Wholesale Electricity Market

As a participant in the two-settlement market, a electricity retailer faces uncertainties from the wholesale market and the real-time consumptions of its customers. If the quantity of consumption is relatively large, the retailer is not a price taker. Instead, its bidding curve and real-time purchase will affect the wholesale price. Using a simplified model, we argue in this section that it is to the retailer's benefit to match the real time consumption with the day-ahead dispatched value. In particular, we motivate, by algebraic and economic arguments, that minimizing the 2-norm deviation of the real-time consumption maximizes the retail surplus. This result motivates the specific form of the cost used in the regret definition in our online learning formulation of the problem.

4.1.1 The day-ahead wholesale market

In this subsection, we consider a simplified model of LMP calculation. In the day-ahead market, the independent system operator (ISO) schedules energy dispatch for the next day. Each electricity generator submits a cost curve $c(p)$ that represents the cost of serving p units of electricity, while each retailer (or Load Serving Entity (LSE)) submits a utility curve $u(d)$ that models the benefit of getting served with d units of electricity. Usually, the day-ahead market dispatch is calculated at the hourly time scale. Therefore, both the demand schedule d and the generation schedule p are 24 dimensional vectors.

With all submitted offers and bids, the ISO solves an optimal power flow (OPF) problem to obtain the optimal dispatch under the objective of maximizing the social welfare. In its simplest form without complications of capacity constrained transmission networks and multiple participating agents, the OPF problem is of the following form,

$$\begin{aligned} \max_{d,p} \quad & u(d) - c(p) \\ \text{s.t.} \quad & d = p \end{aligned} \tag{4.1}$$

The solutions, d^{DA} and p^{DA}, represent the desired day-ahead dispatch of demand and generation. The day-head price is defined as the cost of serving next unit of energy. Therefore, it is the marginal cost of generating p^{DA}, *i.e.*, $\lambda^{\text{DA}} = \frac{\partial c}{\partial p}(p^{\text{DA}})$.

The clearing of the day-ahead market is financially binding in the sense that, regardless of the actual consumption in real time, the day-ahead payment from retailer to the system operator is settled as $(\lambda^{\text{DA}})^{\text{T}} d^{\text{DA}}$. The payment from the system operator to the generator is $(\lambda^{\text{DA}})^{\text{T}} p^{\text{DA}}$. Since the retailer's utility of using d^{DA} is $u(d^{\text{DA}})$, the retail surplus is calculated as,

$$S^{\text{DA}}_{\text{retail}} = u(d^{\text{DA}}) - (\lambda^{\text{DA}})^{\text{T}} d^{\text{DA}}. \tag{4.2}$$

4.1.2 The real-time wholesale market

The actual consumption and generation in real time d^{RT} and p^{RT}, however, are nominally different from the day ahead dispatch. Consequently, the real-time price will deviate from the day-ahead price. In particular, if the cost function of generation in real-time is $\tilde{c}(p)$, the real-time price is calculated as $\lambda^{\text{RT}} = \frac{\partial \tilde{c}}{\partial p}(d^{\text{RT}})$, which stands for the cost of serving the next unit of electricity in real time.

Different from the day-ahead settlement, the real-time settlement only applies to the difference between the day-ahead schedule and the real-time consumption. This means that the payment from the retailer to the system operator is $(\lambda^{\text{RT}})^{\text{T}}(d^{\text{RT}} - d^{\text{DA}})$ if positive. Otherwise, this quantity represents the compensation from the system operator to the retailer.

Therefore, if the real-time consumption matches the day-ahead dispatch, there is no real-time payment. The total retail surplus is still $S^{\text{DA}}_{\text{retail}}$. If the actual consumption d^{RT} is different from d^{DA}, the retail surplus is

$$S^{\text{RT}}_{\text{retail}} = u(d^{\text{RT}}) - [(\lambda^{\text{DA}})^{\text{T}} d^{\text{DA}} + (\lambda^{\text{RT}})^{\text{T}}(d^{\text{RT}} - d^{\text{DA}})], \tag{4.3}$$

where the first term is the utility of the retailer from delivering d^{RT} to its consumer, and the second term is the total payment to the wholesale market. Therefore, the surplus loss due to deviation of d^{RT} from d^{DA} is

$$\Delta S_{\text{retail}} = S^{\text{DA}}_{\text{retail}} - S^{\text{RT}}_{\text{retail}}. \tag{4.4}$$

Based on the Taylor expansion of $u(d^{\text{RT}})$, we can approximate ΔS_{retail} as shown in the following theorem.

Theorem 9 *Under the assumption that cost c(p) and utility u(d) are twice differentiable,*

$$\Delta S_{retail} \approx \theta(d^{RT} - d^{DA})^{T}(d^{RT} - d^{DA}), \tag{4.5}$$

where θ is a constant independent of d^{RT} and d^{DA}

Proof 9 *Consider the first order approximation,*

$$u(d^{RT}) - u(d^{DA}) \approx [\frac{\partial u}{\partial d}(d^{DA})]^{T}(d^{RT} - d^{DA}). \tag{4.6}$$

By the KKT condition and the definition of real-time price

$$\frac{\partial u}{\partial d}(d^{DA}) = \frac{\partial c}{\partial p}(p^{DA}) = \frac{\partial c}{\partial p}(d^{DA}). \tag{4.7}$$

$$\lambda^{RT} = \frac{\partial \tilde{c}}{\partial p}(p^{RT}) = \frac{\partial \tilde{c}}{\partial p}(d^{RT}) = \frac{\partial c}{\partial p}(d^{RT}) + \frac{\partial \Delta c}{\partial p}(d^{RT}), \tag{4.8}$$

where $\Delta c(p) = \tilde{c}(p) - c(p)$. *As the generation cost function,* $c(p)$ *usually takes a quadratic form in practice, i.e., ,* $c(p) = \theta p^T p$, *where* θ *is a scalar. Therefore,*

$$\Delta S_{retail} \approx \theta(d^{RT} - d^{DA})^T(d^{RT} - d^{DA}) + (\frac{\partial \Delta c}{\partial p}(d^{RT}))^T(d^{RT} - d^{DA}). \tag{4.9}$$

Usually, the day-ahead cost function $c(p)$ *and real-time cost function* $\tilde{c}(p)$ *have similar shapes and the perturbation* Δc *has small first order derivative. Hence, compared with the first term in (4.9), the second term can be neglected.*

$$\Delta S_{retail} \approx \theta(d^{RT} - d^{DA})^T(d^{RT} - d^{DA}). \tag{4.10}$$

Therefore, the objective of maximizing retail surplus is equivalent to minimizing the squared deviation of the real-time demand to the day-ahead dispatch.

The result above can also be illustrated in the Price-Quantity plane as shown in Fig. 4.1. The demand function presents the optimal quantity of energy required from the retailer given the price. It is actually the derivative of the utility function $u(d)$. The area below the line is the integration, which is exactly the utility value with quantity d. Similarly, the day-ahead and real-time supply function stand for the optimal quantity of generation to the generator if the price is given. The crossing point (d^{DA}, λ^{DA}) is the day-ahead equilibrium, the same as calculated from (4.1). Subtracting the day-ahead payment from the utility, Area I represents the day-ahead retail surplus.

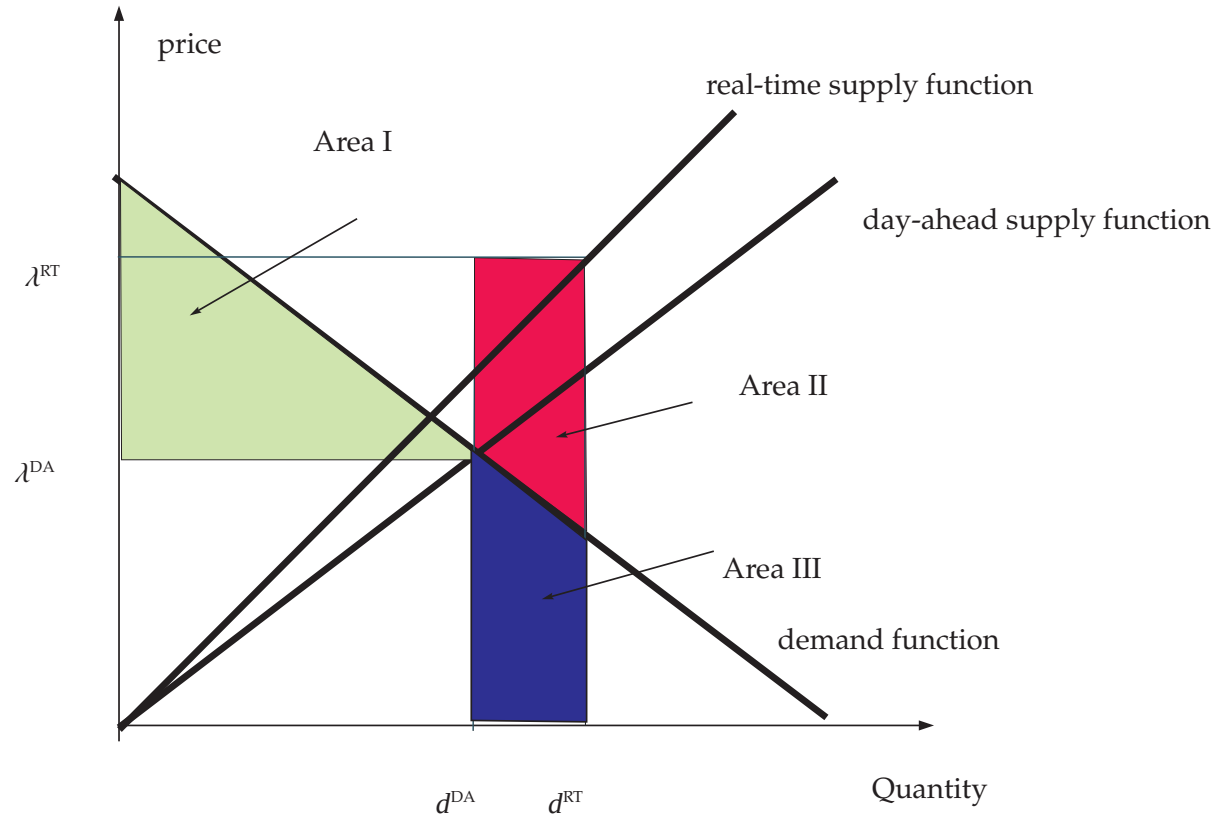


Figure 4.1: Real-time market equilibrium

In the real-time market, the real-time consumption d^{RT} deviates from d^{DA}, and the real-time price, λ^{RT}, is determined by the real-time supply function. Area III is the additional utility gained by consuming d^{RT}, while the sum of area II and III is the real time payment. Therefore, Area II represents the retail surplus loss, and the loss grows in the order of $\|d^{RT} - d^{DA}\|_2^2$—the 2-norm deviation between the day-ahead scheduled consumption and the actual real-time consumption.

4.2 Dynamic Retail Pricing via Online Learning

4.2.1 Pricing policy and regret

As discussed in Section 4.1, minimizing the demand side surplus loss is equivalent, approximately, to minimizing the squared deviation of real-time electricity consumption, d^{RT}, from the day-ahead optimal dispatch, d^{DA}.

Formally, define the t-th day's expected surplus loss as the 2-norm of the deviation of the real-time consumption from the day-ahead dispatch, *i.e.*, $L_t \triangleq \mathbb{E}[\|d_t^{\text{RT}} - d_t^{\text{DA}}\|_2^2]$, where d_t^{DA} and d_t^{RT} are the day-ahead and real-time demands for day t.

Assuming the linear demand function in Theorem 2, for the purpose of obtaining a performance upper bound, we consider the case that the parameters in (3.6), G and b, are known to the retailer. In the following, we will use A instead of G to be consistent with the standard notation in linear regression. At day t, the optimal retail price is given by

$$\pi_t^* = \arg\min_{\pi_t} \mathbb{E}[\|d_t^{\text{RT}} - d_t^{\text{DA}}\|_2^2] = A^{-1}(b - d_t^{\text{DA}}), \tag{4.11}$$

and the corresponding minimum surplus loss is only caused by the exogenous random fluctuations (such as the outdoor temperature). Specifically, the minimized expected loss is

$$\mathbb{E}[\|b - A\pi_t^* + w_t - d_t^{\text{DA}}\|_2^2] = \|\Sigma_w\|_2, \tag{4.12}$$

where Σ_w is the covariance matrix of demand model noise w in (3.6). Notice that the minimized surplus loss is independent of the day-ahead dispatch d_t^{DA}.

However, it is nontrivial for the retailer to obtain the exact parameters of the demand functions of its customers because a customer is likely to consider such information private. At day t, the only information available to the retailer is the record of previous electricity consumption up to $t-1$ and day-ahead dispatch up to t. Formally, the retail pricing policy is defined as follows,

Definition 1 *The retail pricing policy $\mu = (\mu_t)$ is a sequence of mappings where μ_t maps the consumption history and day-ahead demand dispatch to the price vector of day t. In particular, letting π_t^μ be the price vector under policy μ, we have*

$$\pi_t^\mu = \mu_t(d_0^{RT}, ..., d_{t-1}^{RT}, d_0^{DA}, ..., d_{t-1}^{DA}, d_t^{DA}), \tag{4.13}$$

where d_i^{DA}, and d_i^{RT} are the day-ahead dispatch and real-time electricity consumption for day i. □

As for a particular policy μ, the regret R_t^μ at day t is defined as the increase of surplus loss compared with using the optimal price, π_t^*, which means that

$$\begin{aligned} R_t^\mu &\triangleq \mathbb{E}[||b - A\pi_t^\mu + w_t - d_t^{\mathrm{DA}}||_2^2 - ||\Sigma_w||_2] \\ &= \mathbb{E}[||b - A\pi_t^\mu - d_t^{\mathrm{DA}}||_2^2]. \end{aligned} \tag{4.14}$$

Because maximizing the surplus is equivalent to minimizing the regret, we'll focus next on the increasing rate of the cumulative regret up to day T, $\sum_{t=1}^T R_t^\mu$.

4.2.2 Lower bound on the growth rate of regret

To gain insights into the lower bound on the regret, we consider first a simple example when part of the parameters are known. Intuitively, the advantage of knowing partially the parameters should lead to a lower growth rate of regret.

In particular, recall the stochastic affine demand function (3.6) where we assume parameter A is known but b is unknown. Consider the following dynamic pricing policy, $\tilde{\mu}$, given by

$$\pi_t^{\tilde{\mu}} = \bar{\pi}_{t-1} + A^{-1}(\bar{d}_{t-1}^{\mathrm{RT}} - d_t^{\mathrm{DA}}), \tag{4.15}$$

where $\bar{\pi}_{t-1}$ and $\bar{d}_{t-1}$ are the average price and demand up to day $t-1$, *i.e.*,

$$\bar{\pi}_{t-1} = \frac{1}{t}\sum_{i=0}^{t-1}\pi_i \text{ and, } \bar{d}_{t-1}^{\mathrm{RT}} = \frac{1}{t}\sum_{i=0}^{t-1} d_i^{\mathrm{RT}}. \tag{4.16}$$

According to (4.14), straight forward calculation gives that the regret for day t is

$$R_t^{\tilde{\mu}} = \mathbb{E}[||\frac{1}{t}\sum_{i=0}^{t-1} w_i||_2^2] = \frac{1}{t}||\Sigma_w||_2. \tag{4.17}$$

Therefore, the aggregated regret,

$$\sum_{t=1}^{T} R_t^{\tilde{\mu}} = \sum_{t=1}^{T}\frac{1}{t}||\Sigma_w||_2 \leq (1+\log T)||\Sigma_w||_2 \tag{4.18}$$

Therefore, with the knowledge of the demand function parameter A, the policy $\tilde{\mu}$ achieves the aggregated regret $O(\log T)$ for any b and any arbitrary sequence of $\{d_t^{\mathrm{DA}}\}$.

To establish the actual lower bound on the growth rate of regret, we formulate a game that, after the retailer proposes a deterministic pricing policy μ, there exists an adversary designing parameters of the demand function. The adversary is to create the worst loss to the retailer while the retailer tries to minimize the largest possible loss. In other words, we consider the following the min-max regret as the objective,

$$\min_{\mu}\max_{b,A}\sum_{t=1}^{T} R_t^{\mu}.$$

The following theorem shows that in the min-max sense, the growing rate of the cumulative regret can not be lower than $\log T$.

Theorem 10 *For any pricing policy μ as defined in (4.13), there exist some $(A, b, d_0^{DA}, ..., d_{t-1}^{DA}, d_t^{DA}, ...)$ to make the cumulative regret, $\sum_{t=1}^{T} R_t^{\mu}$, grows at least at the rate of* $\log T$.

Proof 10 *First, we reduce the problem to the case that A is known, d_i^{DA}'s are constant and Σ_w is a diagonal matrix with the diagonal elements all as σ_w^2. The minimax rate for this case lower bounds the general case.*

For any policy μ, the maximum regret among all possible b is

$$L(\mu) = \max_{b} \sum_{t=1}^{T} R_t^{\mu}. \tag{4.19}$$

Assume the parameter b follows a prior distribution,γ_n : $\mathcal{N}(\bar{b}, n\sigma^2 I)$, where n is a positive integer number and I is an identity matrix. Define the Bayesian cost as (4.14) and denote the Bayesian estimator of b as η_n. By the property of joint Gaussian distribution and Sherman-Morrison formula, we can get the minimum Bayesian risk,

$$R_t^{\eta_n}(\gamma_n) = \mathbb{E}^{\eta_n}\|b - \eta_n(d_1^{RT}, ..., d_{t-1}^{RT})\|_2^2 = \frac{n\sigma^2}{\sigma_w^2 + tn\sigma^2}\|\Sigma_w\|_2$$

Then, the cumulative Bayesian risk $\sum_{t=1}^{T} R_t^{Bayes}(\gamma_n)$ is an increasing function of n and goes to $L(\tilde{\mu}) = \sum_{t=1}^{T} \frac{1}{t}\|\Sigma_w\|_2$ as n goes to ∞, where $\tilde{\mu}$ is defined in Eq. (4.15).

If $\tilde{\mu}$ is not the minimax estimator, there exist some policy $\bar{\mu}$ and $\epsilon > 0$, s.t. $L(\bar{\mu}) < L(\tilde{\mu}) - \epsilon$. On the other hand, for $\epsilon > 0$, we can find some positive integer m, s.t.

$$L(\tilde{\mu}) - \epsilon < \sum_{t=1}^{T} R_t^{\eta_m}(\gamma_m).$$

By the definition (4.19), the Bayesian risk of policy $\bar{\mu}$ under the distribution γ_m should be less than the maximum cost over all possible values of b, i.e.,

$$\sum_{t=1}^{T} R_t^{\bar{\mu}}(\gamma_m) < L(\bar{\mu}) < L(\tilde{\mu}) - \epsilon < \sum_{t=1}^{T} R_t^{\eta_m}(\gamma_m),$$

which contradicts the fact that η_m is the Bayesian estimator.

4.2.3 PWLSA: a rate optimal learning policy

In this section, we propose a policy that achieves the lower bound on the regret growth rate; it is thus optimal in the sense of having the lowest rate of growth. Referred to as piecewise linear stochastic approximation (PWLSA) policy, the proposed policy is an extension of the stochastic approximation approach of Lai and Robin [41] for scaler processes with a single desired optimal price.

If the day-ahead demand is the same for all days, stochastic approximation will use the previous average price as the nominal value and previous average demand as the feedback signal to calculate the next price, as shown below,

$$\pi_t^{\text{SA}} = \overline{\pi}_{t-1} + \gamma(\overline{d}_{t-1}^{\text{RT}} - d^{\text{DA}}), \tag{4.20}$$

where d^{DA} is the constant day-ahead dispatch level, and the feedback factor γ is a positive scalar.

For multiple day-ahead dispatch levels, we build adaptively a dictionary of day-ahead dispatch levels that have appeared before. Denote the dictionary at day t as $\mathcal{D}_t$. For day $t+1$, if $d_{t+1}^{\text{DA}} \in \mathcal{D}_t$, let $\mathcal{D}_{t+1} = \mathcal{D}_t$. Otherwise, $\mathcal{D}_{t+1} = \mathcal{D}_t \bigcup \{d_{t+1}^{\text{DA}}\}$. For each day-ahead dispatch level in $\mathcal{D} = \bigcup_{t=1}^{\infty} \mathcal{D}_t$, we keep a separate stochastic approximation to calculate the retail price, in a feedback control fashion similar as (4.20).

Therefore, for different d_t^{DA}, we have a different linear function to calculate the next retail price. The policy is piecewise linear. Formally, the PWLSA policy, μ^{PWLSA}, is defined as,

Definition 2 (PWLSA) *Assume for all $t \in \mathbb{N}^+$, $d_t^{DA} \in \mathcal{D}$ and $\mathcal{D}$ is countable.*

- *If* $d_t^{DA} \in \mathcal{D}_t$, *then* $\mathcal{D}_{t+1} = \mathcal{D}_t$ *and*

$$\pi_t^{PWLSA} = \frac{1}{|\mathcal{C}_t^{d_t^{DA}}|}\left(\sum_{k \in \mathcal{C}_{d_t^{DA}}} \pi_k^{PWLSA} + \gamma(d_k^{PWLSA} - d_t^{DA})\right), \tag{4.21}$$

where $\mathcal{C}_t^{d_t^{DA}} = \{k \in \mathbb{N}^+ : k \leq t-1, d_k^{DA} = d_t^{DA}\}$ *and* $|\mathcal{C}_t^{d_t^{DA}}|$ *is the total number of elements in* $\mathcal{C}_t^{d_t^{DA}}$.

- *Otherwise,* $d_t^{DA} \notin \mathcal{D}_t$, *then* $\mathcal{D}_{t+1} = \mathcal{D}_t \bigcup \{d_t^{DA}\}$ *and*

$$\pi_t^{PWLSA} = \tilde{\pi}_j, \tag{4.22}$$

where $\tilde{\pi}_j$ *is an arbitrary predetermined price.*

□

The following theorem shows that PWLSA can achieve the optimal logarithmic regret order.

Theorem 11 *Assume that day-ahead dispatch* d_t^{DA}*'s are from a finite set, i.e.,* $|\mathcal{D}| < \infty$. *If* $\gamma \geq \frac{1}{2\lambda_{\min}(A)}$, *where* $\lambda_{\min}(A)$ *is the minimum eigenvalue of A, then we have,*

$$\sum_{t=1}^{T} R_t^{\mu^{PWLSA}} \sim O(\log(T)), \tag{4.23}$$

Proof 11 *First, we consider when there is a single day-ahead dispatch level* d^{DA}, *and* $\pi^* = A^{-1}(b - d^{DA})$. *After simplification,*

$$\begin{aligned} &\pi_{n+1} - \pi_{n+1}^* \\ &= (I - \gamma A)[\Pi_{i=1}^{n-1}(1 - \tfrac{\gamma A}{i+1})](\pi_1 - \pi^*) \\ &\quad + \textstyle\sum_{k=1}^{n}\{\frac{\gamma}{n} + \sum_{j=k}^{n-1}[\Pi_{i=j+1}^{n-1}(I - \frac{\gamma A}{i+1})]\frac{(I-\gamma A)\gamma}{j(j+1)}\}\omega_k. \end{aligned} \tag{4.24}$$

For the first term in (4.24),

$$\|(I - \gamma A)[\Pi_{i=1}^{n-1}(1 - \frac{\gamma A}{i+1})]\|_2^2 \leq \|(I - \gamma A)\|_2^2 \Pi_{i=1}^{n-1}\|(I - \frac{\gamma A}{i+1})\|_2^2.$$

Since $(I - \frac{\gamma A}{i+1})^T(1 - \frac{\gamma A}{i+1}) = I - \frac{2\gamma A}{i+1} + \frac{\gamma^2 A^2}{(i+1)^2}$, *denoting* λ_m *as the minimum eigenvalue of* A, *we have,*

$$\|(I - \frac{\gamma A}{i+1})\|_2^2 \leq I - \frac{2\gamma\lambda_m}{i+1} + \frac{\gamma^2}{(i+1)^2}\|A\|_2^2.$$

Let $C_1 \triangleq \|I - \gamma A\|_2^2$. *Then, since* $\gamma\lambda_m > \frac{1}{2}$

$$\begin{aligned}&\|(I - \gamma A)[\Pi_{i=1}^{n-1}(I - \tfrac{\gamma A}{i+1})]\|_2^2\\ &\leq C_1\Pi_{i=1}^{n-1}(I - \tfrac{2\gamma\lambda_m}{i+1} + \tfrac{\gamma^2}{(i+1)^2}\|A\|_2^2) = C_2\tfrac{1}{n+1},\end{aligned}$$

where $C_2 = C_1 exp\{\gamma^2\|A\|_2^2\}$ *doesn't depend on* n.

For the second term in (4.24),

$$\begin{aligned}&\|[\Pi_{i=j+1}^{n-1}(I - \tfrac{\gamma A}{i+1})](I - \gamma A)\|_2^2\\ &\leq C_1 exp\{\textstyle\sum_{i=j+1}^{n} -\frac{2\gamma\lambda_m}{i+1} + \frac{\gamma^2}{(i+1)^2}\|A\|_2^2\} \leq C_2(\tfrac{j+1}{n+1})^{2\gamma\lambda_m}.\end{aligned}$$

Then,

$$\begin{aligned}&\|\tfrac{\gamma}{n} + \textstyle\sum_{j=k}^{n-1}[\Pi_{i=j+1}^{n-1}(I - \frac{\gamma A}{i+1})]\frac{(1-\gamma A)\gamma}{j(j+1)}\|_2^2\\ &\leq \{\tfrac{\gamma}{n} + \textstyle\sum_{j=k}^{n-1}\|[\Pi_{i=j+1}^{n-1}(I - \frac{\gamma A}{i+1})](I - \gamma A)\|_2\frac{\gamma}{j(j+1)}\}^2\\ &\leq 2\tfrac{\gamma^2}{n^2} + 2\gamma C_2(\tfrac{1}{n})(\tfrac{1}{n})^{2\gamma\lambda_m - 1}(\tfrac{1}{k})^{2-2\gamma\lambda_m}.\end{aligned}$$

Sum the two terms up,

$$\begin{aligned}&\textstyle\sum_{k=1}^{n-1}\|\frac{\gamma}{n} + \sum_{j=k}^{n-1}[\Pi_{i=j+1}^{n-1}(I - \frac{\gamma A}{i+1})]\frac{(1-\gamma A)\gamma}{j(j+1)}\|_2^2\\ &\leq 2\tfrac{\gamma^2}{n} + 2\gamma C_2(\tfrac{1}{n})(\tfrac{1}{n})^{2\gamma\lambda_m - 1}\textstyle\sum_{k=1}^{n-1}(\frac{1}{k})^{2-2\gamma\lambda_m} \leq C_3\frac{1}{n}.\end{aligned}$$

Define $M = \max\{\|\pi_1 - \pi^*\|_2^2, \|\Sigma_\omega\|_2^2, \|\Sigma_d\|_2^2\}$, *we have*

$$\begin{aligned}\textstyle\sum_{i=1}^{n} L_n &= \mathbb{E}\textstyle\sum_{i=1}^{n}\|A(\pi_i - \pi^*)\|_2^2\\ &\leq \textstyle\sum_{n=1}^{T}\|A\|_2^2[(C_2 + C_1)\frac{1}{n}]M \leq C\log(T).\end{aligned}$$

If $|\mathcal{D}|$ *is finite, and we use a separate stochastic approximation to calculate the retail price, then the accumulated regret* $\sum_{n=1}^{T} R_n \leq C|\mathcal{D}|\log(T)$.

Since $\log T$ is shown to be the optimal rate achievable, PWLSA is already the best in the sense of asymptotic growing rate of the regret. The conditions in the theorem are quite general. In practice, the cost functions from the generator and the utility functions from the retailer won't change often and are usually chosen from a few alternatives. Therefore, we can assume the total number of possible day-ahead dispatch levels, $|\mathcal{D}|$, to be finite. On the other hand, the consumers' demand function is from real data, which can be constrained by a compact set. Therefore, the bound of the minimum eigenvalue of A is not hard to get with reasonable assumption.

4.3 Numerical results

4.3.1 Simulation set-up

In this section, we conducted simulations based on the actual temperature records in Hartford, CT, from July 1st, 2012 to July 30th, 2012. The day-head price was also for the same period from ISO New England. The HVAC parameters for the simulation were set as: $\alpha = 0.5, \beta = 1, \mu = 10$. The desired indoor temperature was set to be $18^{\circ}C$ for all hours. The size of aggregation was assumed to be 100.

4.3.2 Learning static parameters

First, we examined PWLSA's ability to identify the correct price if the parameters of the demand model remain the same. To make the comparison, we used

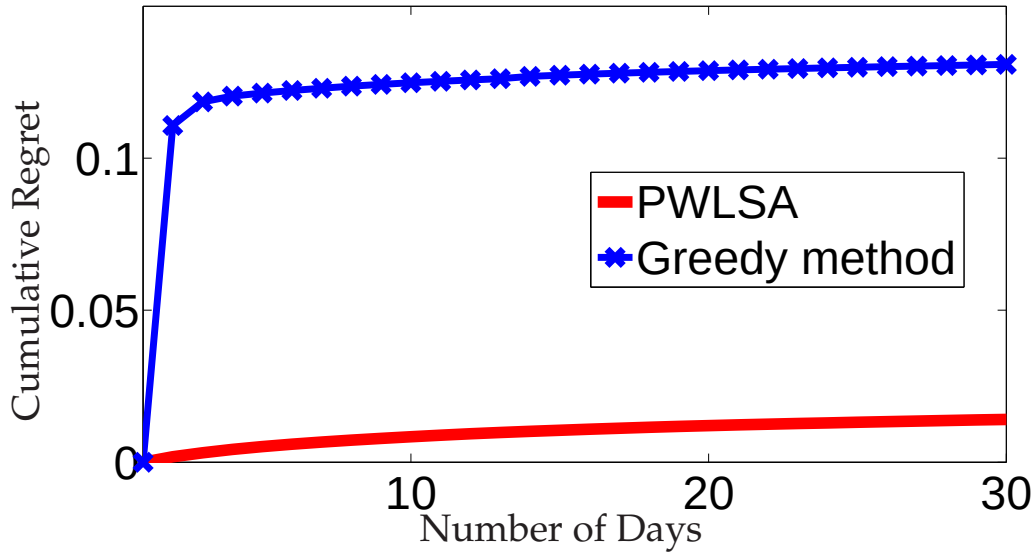


Figure 4.2: Cumulative regret

the Greedy Method [7, 47] as a benchmark. At each day, the Greedy Method makes maximum likelihood estimate of the parameters and uses the result as the correct parameters to calculate the "optimal" price.

Fig. 4.2 and Fig. 4.3 show the average performance of PWLSA and Greedy Method over 10,000 Monte Carlo runs. In Fig. 4.2, the induced cumulative regrets of the two policies are compared. We could identify the logarithmic growth of the cumulative regret under PWLSA and significant cumulative regret increase by the Greedy Method. Fig. 4.3 shows the absolute percentage deviation of the prices under the two polices from the optimal price. We can see that Greedy Method performed extremely bad at the very beginning due to insufficient learning. After some days, the two policies both produced prices pretty close to the optimal one.

After carefully investigating the simulated data, we found two typical scenarios as shown in Fig. 4.4 and Fig. 4.5. We used the ratio of the calculated price to the optimal price as y-axis, to show the fluctuation. In most of the cases as

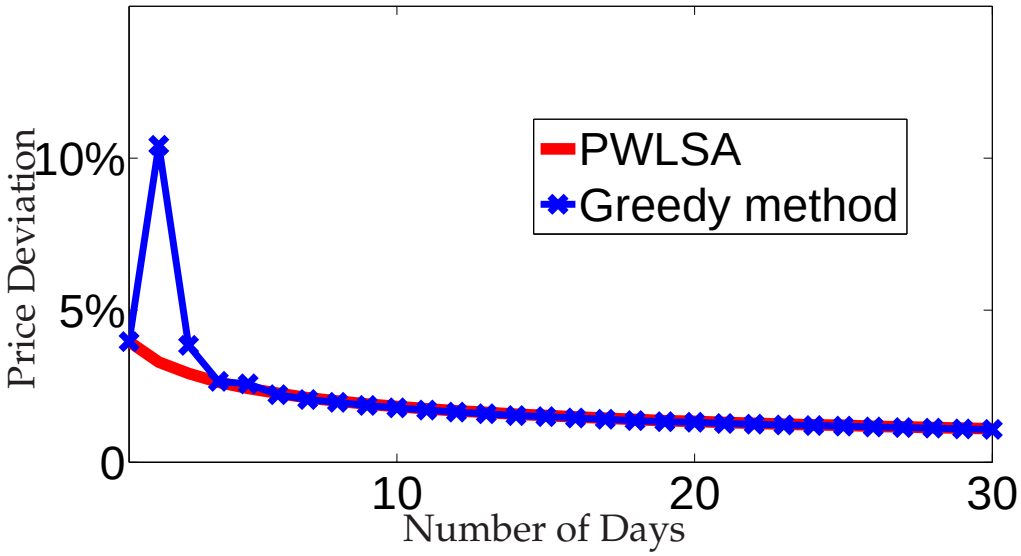


Figure 4.3: Price convergence

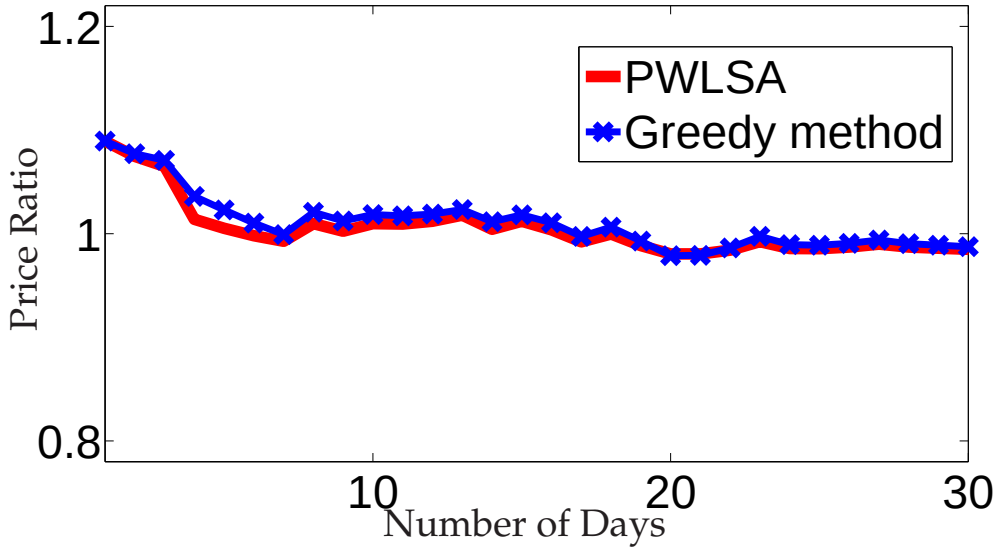


Figure 4.4: Scenario 1: The two had similar performance

in Fig. 4.4, the two polices gave similar performance and both converged to the optimal price fast. On the other hand, Fig. 4.5 shows one extreme scenario that Greedy Method run into the condition that is close to singularity, which leads to an abnormal price. Although this kind of scenarios happened rarely, it caused the wide performance gap between the Greedy Method and PWLSA.

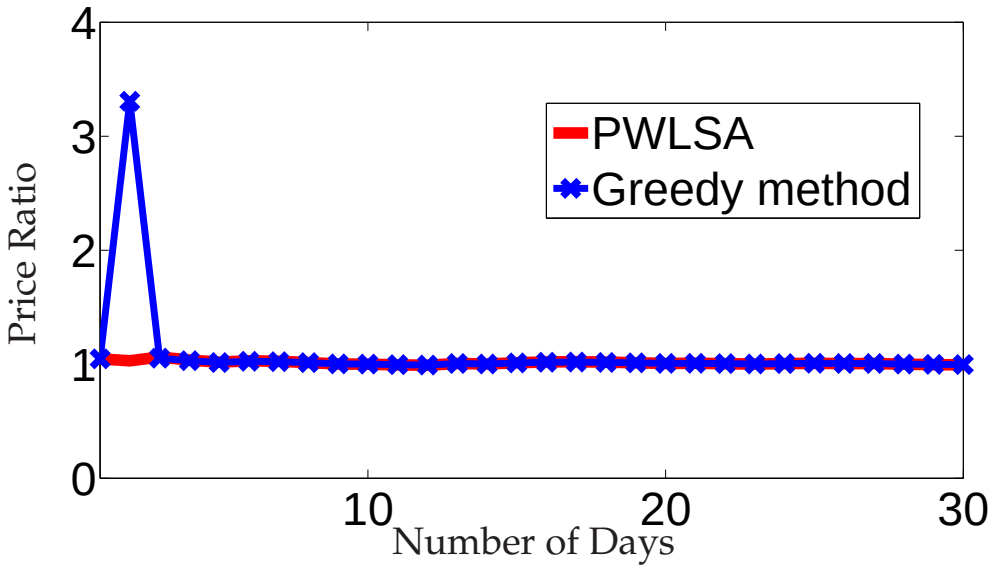


Figure 4.5: Scenario 2: the Greedy Method failed

4.3.3 Learning dynamic parameters

In the real world, the parameters of the demand model usually do not stay constant. They may follow some cycles or drifts. In this subsection, we tested the learning ability and robustness of PWLSA under dynamic unknown parameters. Besides the set of parameters above, we used $1.5A$ instead of A to make the alternative set of parameters. We assumed the parameters followed a Markov Chain with these two sets as states. The transition probability to the other set was assumed to be 0.25.

Fig. 4.6 and Fig. 4.7 show the average performance comparison of PWLSA and the Greedy Method under dynamic unknown demand model. We can see that PWLSA still outperformed the Greedy Method. According to Fig. 4.6, the cumulative regret under PWLSA grew linearly. Intuitively, when a sequence of observation is given, a policy will produce a fixed price or a fixed probability distribution over candidate prices (for randomized policy). However, since

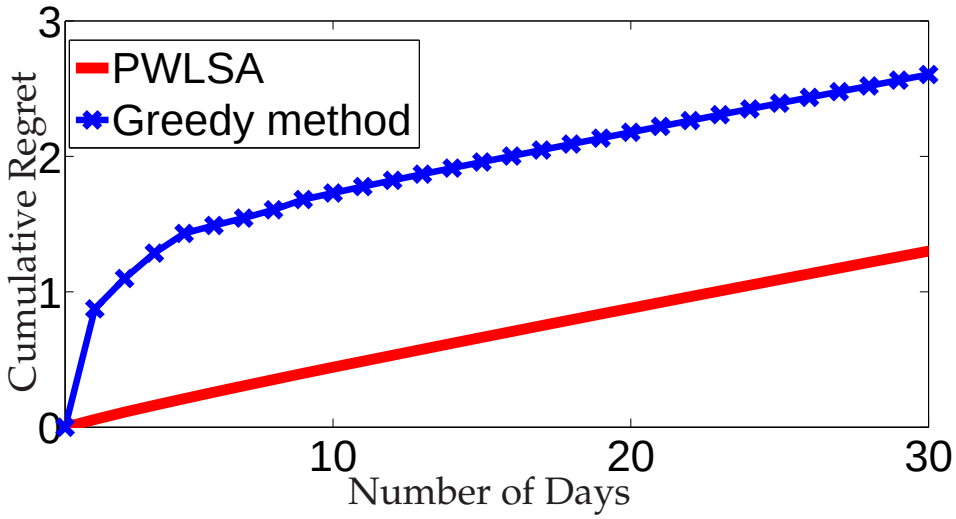


Figure 4.6: Cumulative regret

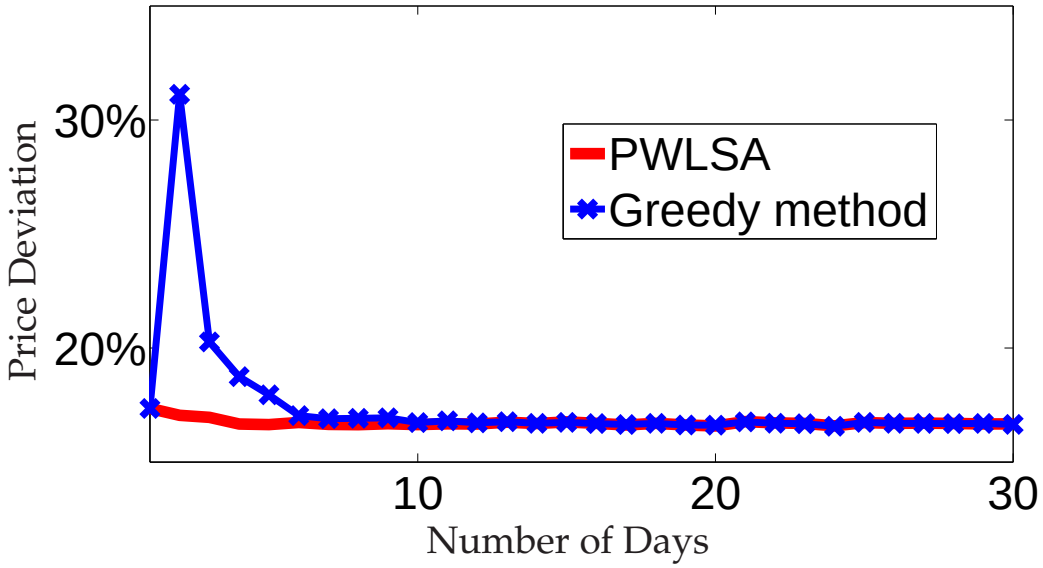


Figure 4.7: Price convergence

the next optimal price is random, there always exists a fixed addition to the expected cumulative regret. Therefore, the linear order achieved by PWLSA is already the best. Fig. 4.7 shows that the error at the very beginning was the cause of the performance gap between PWLSA and the Greedy Method.

We also conducted the scenario analysis similar to the static parameter case as shown in Fig. 4.8 and Fig. 4.9. Each changing point stands for a incident that

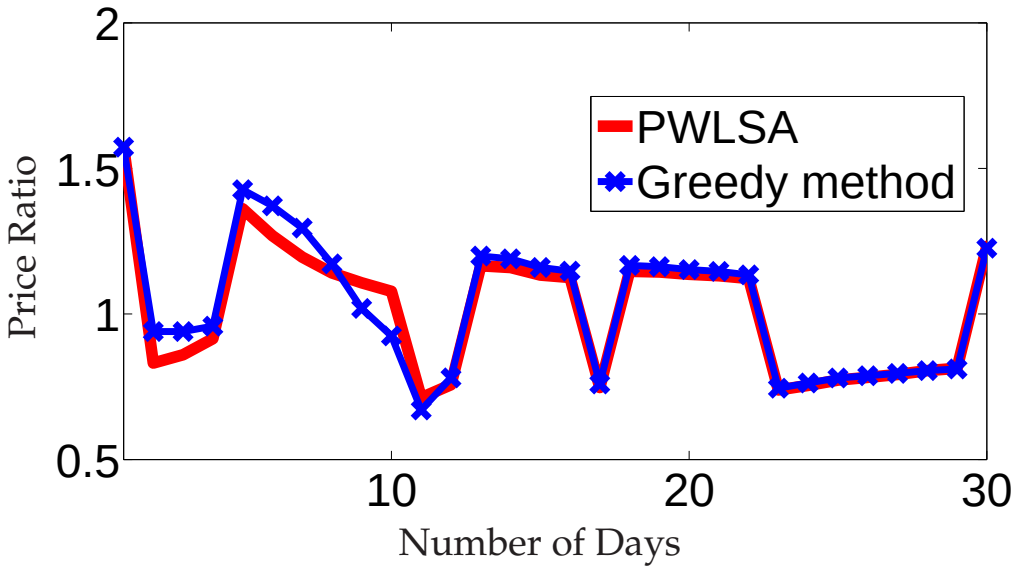


Figure 4.8: Scenario 1: The two had similar performance

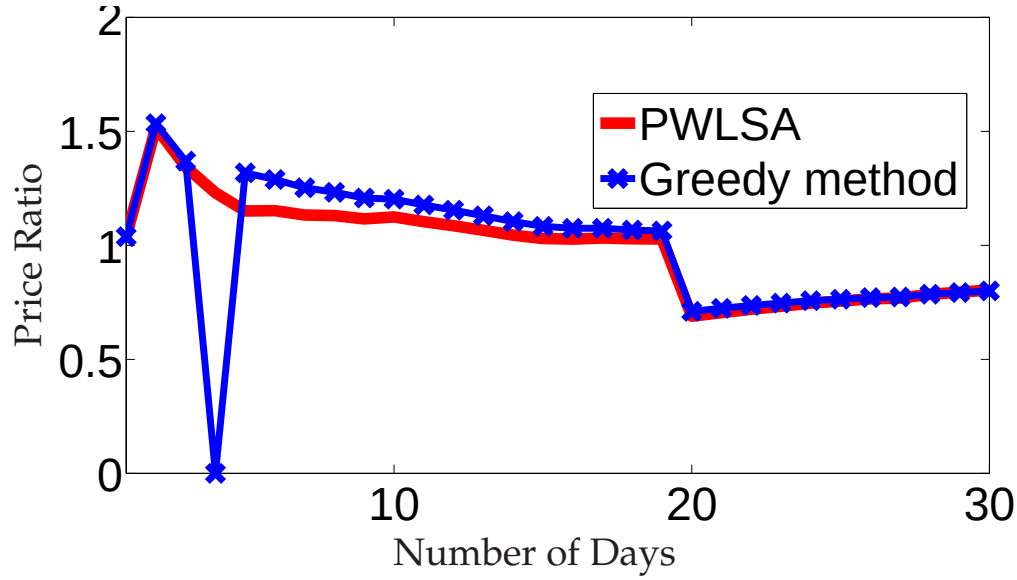


Figure 4.9: Scenario 2: the Greedy Method failed

the state jumps to the other set. In Fig. 4.8, we can see that the two policies had similar performance and both tracked the optimal prices well. In few extreme cases, the Greedy Method lost track on the optimal prices wildly at the beginning a few days, as shown in Fig. 4.9.

CHAPTER 5

CONCLUSIONS

In this thesis, we focus on three topics related to real-time electricity pricing: the effect of data quality on real-time wholesale price, the interaction between electricity retailer and its consumers, and the online learning of dynamic retail price.

5.1 Effect of data quality on real-time wholesale price

We report in this part a study on impacts of worst data on the real-time market operation. A key result is the geometric characterization of real-time LMP given in Theorem 1. This result provides insights into the relation between data and the real-time LMP; it serves as the basis of characterizing impacts of bad data.

Our investigation includes bad data scenarios that arise from both analog meter measurements and digital breaker state data. To this end, we have presented a systematic approach by casting the problem as one involving an adversary injecting malicious data. While such an approach often gives overly conservative analysis, it can be used as a measure of assurance when the impacts based on worst case analysis are deemed acceptable.

We note that, because we use adversary attacks as a way to study the worst data, our results have direct implications when cyber-security of smart grid is considered. Given the increasing reliance on information networks, developing effective countermeasures against malicious data attack on the operations of a future smart grid is crucial. See [38, 35, 27, 34] for discussion about countermeasures.

From a practical viewpoint, our result can serve as the guideline to the real-time operation. Following the methodology in this thesis, worst effect of a specific set of meters on real-time LMP can be checked. Once a huge potential perturbation is detected, alarm should be made and the operator needs to check the accuracy of these specific data, add protection devices, or even add more redundant meters.

Although our findings are obtained from academic benchmarks involving relatively small size networks, we believe that the general trend that characterizes the effects of bad data is likely to persist in practical networks of much larger size. In particular, as the network size increases and the number of simultaneous appearance of bad data is limited, the effects of the worst meter data on LMP decrease whereas the effects of the worst topology data stay nonnegligible regardless of the network size. This observation suggests that the bad topology data are potentially more detrimental to the real-time market operation than the bad meter data.

5.2 Interactions between electricity retailer and consumers

In this part, we study a day ahead hourly pricing (DAHP) mechanism for distributed demand response in uncertain and dynamic environments. Such a pricing scheme has the advantage of reducing consumer anxiety of pricing uncertainties and allowing the retailer to optimize the retail pricing adaptively.

We formulate the problem as a Stackelberg game in which the retailer plays the role of a leader and the consumers the followers. For thermal dynamic load, we obtain an affine form of optimal demand response to DAHP from which the

retailer optimizes payoff functions of its own and characterize a concave trade-off between consumer surplus (CS) and retail profit (RP).

The DAHP framework also provides insights into the role of renewable sources and storage devices. We show that for either renewable energy or storage devices, the effects on CS-RP trade-off are significantly different depending on whether they are on the retailer side or the consumer side.

5.3 Online learning of dynamic retail price

We present in this part an online learning approach to the dynamic pricing of electricity of a retailer whose customers have price responsive dynamic load with unknown demand function. We exploit the linear form of the demand function for thermal dynamic load, and cast the problem of online learning as tracking day-ahead dispatch. This approach leads to a simple learning algorithm with the growth rate of cumulative regret at the order of $\log T$, which is the best rate achievable for any dynamic pricing policies.

BIBLIOGRAPHY

[1] GTM Research, U.S. SOLAR MARKET INSIGHT, 2014.

[2] Residential Energy Consumption Survey (RECS), http://www.eia.gov/consumption/residential/.

[3] National Solar Radiation Data Base, http://rredc.nrel.gov/solar/old_data/nsrdb.

[4] Power systems test case archive. `http://www.ee.washington.edu/research/ps`

[5] A. Abur and A. G. Expósito. *Power system state estimation: theory and implementation*. CRC, 2000.

[6] R. Agrawal. The continuum-armed bandit problem. *SIAM Journal on Control and Optimization*, 33(6):1926–1951, 1995.

[7] T. W. Anderson and J. Taylor. Some experimental results on the statistical properties of least squares estimates in control prob- lems. *Econometrica*, 44:12891302, 1976.

[8] P. Auer, R. Ortner, and C. Szepesvari. Improved rates for the stochastic continuum-armed bandit problem. *Lecture Notes in Computer Science*, 4539:454–468, 2007.

[9] D. Bargiotas and J. Birddwell. Residential air conditioner dynamic model for direct load control. *IEEE Transactions on Power Delivery*, 3(4):2119–2126, Oct. 1988.

[10] D. Bertsimas and G. Perakis. Dynamic pricing: a learning approach. *Mathematical and Computational Models for Congestion Charging*, pages 45–80, 2006.

[11] S. Borenstein. The long run efficiency of real-time electricity pricing. *The Energy Journal*, 26(3), 2005.

[12] S. Borenstein, M. Jaske, and A. Rosenfeld. Dynamic pricing, advanced metering, and demand response in electricity markets. *Recent Work, Center for the Study of Energy Markets, UC Berkeley*, 2002.

[13] J. Broder and P. Rusmevichientong. Dynamic pricing under a general parametric choice model. *Operations Research*, 60(4):965–980, 2012.

[14] D. Cai, S. Adlakha, S. Low, P. de Martini, and K. Chandy. Impact of residential pv adoption on retail electricity rates. *Energy Policy*, July 2013.

[15] M. Carrion, A. Conejo, and J. Arroyo. Forward contracting and selling price determination for a retailer. *IEEE Transactions on Power Systems*, 32(4), Nov. 2007.

[16] K. A. Clements and P. W. Davis. Detection and identification of topology errors in electric power systems. *IEEE Transactions on Power Systems*, 3(4):1748 –1753, Nov. 1988.

[17] A. Conejo, R. Garcia-Bertrand, M. Carrion, A. Caballero, and A. Andres. Optimal involvement in futures markets of a power producer. *IEEE Transactions on Power Systems*, 23(2), May 2008.

[18] E. W. Cope. Regret and convergence bounds for a class of continuum-armed bandit problems. *IEEE Transactions on Automatic Control*, 54(6), Jan. 2009.

[19] I. S. Costa and J. A. Leao. Identification of topology errors in power system state estimation. *IEEE Transactions on Power Systems*, 8(4):1531 –1538, Nov. 1993.

[20] FERC. Assessment of demand response and advanced metering, series 01-12, Dec. 2008.

[21] A. Garcia, E. Campos-Nanez, and J. Reitzes. Dynamic pricing and learning in electricity markets. *Operation Research*, 53(2):231–241, 2005.

[22] R. Gibbons. *Game Theory for Applied Economists*. Princeton University Press, 1992.

[23] E. Handschin, F. C. Schweppe, J. Kohlas, and A. Fiechter. Bad data analysis for power system state estimation. *IEEE Transactions on Power Apparatus and Systems*, PAS-94(2):329–337, Mar./Apr. 1975.

[24] J. M. Harrison, N. B. Keskin, and A. Zeevi. Bayesian dynamic pricing policies: learning and earning under a binary prior distribution. *Management Science*, pages 1–17, Oct. 2011.

[25] N. Hopper, C. Goldman, and B. Neenan. demand response from day-ahead hourly pricing for large customers. *Electricity Journal*, (02):52–63, 2006.

[26] S. Huang, Q. Wu, S.S. Oren, R. Li, and Z. Liu. Distribution locational marginal pricing through quadratic programming for congestion management in distribution networks. *IEEE Transactions on Power Systems*, PP(99):1–9, 2014.

[27] G. Hug and J. A. Giampapa. Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks. *IEEE Transactions on Smart Grid*, 3(3):1362–1370, 2012.

[28] L. Jia, R. J. Thomas, and L. Tong. On the nonlinearity effects on malicious data attack on power system. In *2012 Power and Energy Society general meeting*, July 2012.

[29] L. Jia and L. Tong. Optimal pricing for residential demand response: a stochastic optimization approach. In *2012 Allerton Conference on Communication, Control and Computing*, Oct. 2012.

[30] L. Jia and L. Tong. Day ahead dynamic pricing for demand response in dynamic environments. In *52nd IEEE Conference on Decision and Control*, Dec. 2013.

[31] J.-Y. Joo and M. Ilic. Multi-layered optimization Of demand resources using Lagrange dual decomposition. *IEEE Transactions on Smart Grid*, 2013.

[32] Andrew Keane and Mark OMalley. Optimal Allocation of Embedded Generation on Distribution Networks. *IEEE TRANSACTIONS ON POWER SYSTEMS*, 20(3), Aug. 2005.

[33] J. Kim and L. Tong. On topology attack of a smart grid. In *2013 IEEE PES Innovative Smart Grid Technologies (ISGT)*, Washington, DC, Feb. 2013.

[34] J. Kim and L. Tong. On topology attack of a smart grid: undetectable attacks and countermeasures. *IEEE Journal on Selected Areas in Communications*, 31(7), July 2013.

[35] T.T. Kim and H.V. Poor. Strategic protection against data injection attacks on power grids. *IEEE Transactions on Smart Grid*, 2(2):326 –333, June 2011.

[36] R. Kleinberg. Nearly tight bounds for the continuum-armed bandit problem. *Advances in Neural Information Processing Systems*, pages 697–740, 2004.

[37] R. Kleinberg and T. Leighton. The value of knowing a demand curve:

bounds on regret for online posted-price auctions. In *Proc. 44th IEEE Symposium on Foundations of Computer Science (FOCS)*, 2003.

[38] O. Kosut, L. Jia, R. J. Thomas, and L. Tong. Malicious data attacks on the smart grid. *IEEE Transactions on Smart Grid*, 2(4):645 –658, dec. 2011.

[39] I. Koutsopoulos, V. Hatzi, and L. Tassiulas. Optimal energy storage control policies for the smart power grid. In *Proc. of IEEE International Conference on Smart Grid Communications*, pages 475–480, 2011.

[40] T. L. Lai and H. Robbins. Adaptive design and stochastic approximation. *The Annals of Statistics*, 7(6):1196–1221, 1979.

[41] T. L. Lai and H. Robbins. Iterated least squares in multiperiod control. *Advanced and Applied Mathematics*, 3:50–73, 1982.

[42] T. L. Lai and H. Robbins. Asymptotically efficient adaptive allocation rules. *Advanced and Applied Mathematics*, 6(1):4–22, 1985.

[43] N. Li, L. Chen, and S. H. Low. Optimal demand response based on utility maximization in power networks. In *IEEE Power and Energy Society General Meeting*, 2011.

[44] Na Li, Lijun Chen, and Steven H. Low. Demand Response in Radial Distribution Networks: Distributed Algorithm. In *Asilomar Conference*, 2012.

[45] E. Litvinov, T. Zheng, G. Rosenwald, and P. Shamsollahi. Marginal loss modeling in lmp calculation. *IEEE Transaction on Power System*, 19(2), 2004.

[46] Y. Liu, P. Ning, and M. K. Reiter. False data injection attacks against state estimation in electric power grids. In *ACM Conference on Computer and Communications Security*, pages 21–32, 2009.

[47] M. Lobo and S. Boyd. Pricing and learning with uncertain demand. In *INFORMS Revenue Management Conference*, Columbia University, 2003.

[48] P. Luh, Y. C. Ho, and R. Muralidharan. Load adaptive pricing: an emerging tool for electric utilities. *IEEE Transactions on Automatic Control*, AC-27(2), Apr. 1982.

[49] A. Mas-Colell and M. D. Whinston. *Microeconomics Theory*. Oxford University Press, 1995.

[50] A. Monticelli. Modeling circuit breakers in weighted least squares state estimation. *IEEE Transactions on Power Systems*, 8(3):1143 –1149, Aug. 1993.

[51] D. O'Neill, M. Levorato, A. Goldsmith, and U. Mitra. Residential demand response using reinforcement learning. In *Proceedings of 2010 First IEEE International Conference on Smart Grid Communications*, pages 409–414, Oct. 2010.

[52] A. L. Ott. Experience with pjm market operation, system design, and implementation. *IEEE Transactions on Power Systems*, 18(2):528–534, May 2003.

[53] A. Papavasiliou and S.S. Oren. Coupling wind generators with deferrable loads. In *IEEE Energy2030*, pages 1–7, Nov. 2008.

[54] A. Papavasiliou and S.S. Oren. Integrating renewable energy contracts and wholesale dynamic pricing to serve aggregate flexible loads. In *IEEE Power and Energy Society General Meeting*, 2011.

[55] T. Pinto, Z. Vale, F. Rodrigues, and I. Praca. Cost dependent strategy for electricity markets bidding based on adaptive reinforcement learning. In *The 16th International Conference on Intelligent System Application to Power Systems*, 2011.

[56] Z. Qiu, E. Peeters, and G. Deconinck. Comparison of two learning algorithms in modelling the generator's learning abilities. In *The 15th International Conference on Intelligent System Applications to Power Systems*, 2009.

[57] A. Rahimi-Kian, B. Sadeghi, and R. J. Thomas. Q-learning based supplier-agents for electricity markets. In *IEEE Power Engineering Society General Meeting*, 2005.

[58] P. Rusmevichientong and J. N. Tsitsiklis. Linearly parameterized bandits. *Mathematics of Operations Research*, 35(2):395–411, 2010.

[59] Thipnatee Sansawatt, Luis F. Ochoa, and Gareth P. Harrison. Smart Decentralized Control of DG for Voltage and Thermal Constraint Management. *IEEE TRANSACTIONS ON POWER SYSTEMS*, 27(3), 2012.

[60] F. C. Schweppe, J. Wildes, and D. P. Rom. Power system static state estimation, Parts I, II, III. *IEEE Transactions on Power Apparatus and Systems*, PAS-89:120–135, 1970.

[61] Mostafa F. Shaaban, Yasser M. Atwa, and Ehab F. El-Saadany. DG Allocation for Benefit Maximization in Distribution Networks. *IEEE TRANSACTIONS ON POWER SYSTEMS*, 28(2), May 2013.

[62] H. I. Su and A. El Gamal. Modeling and analysis of the role of fast-response energy storage in the smart grid. In *Proc. of the 49th Annual Allerton Conference*, Oct. 2011.

[63] Hideharu Sugihara, Kohei Yokoyama, Osamu Saeki, Kiichiro Tsuji, and Tsuyoshi Funaki. Economic and Efficient Voltage Management Using Customer-Owned Energy Storage Systems in a Distribution Network With High Penetration of Photovoltaic Systems . *IEEE TRANSACTIONS ON POWER SYSTEMS*, 28(1), Feb. 2013.

[64] Sicong Tan, Jian-Xin Xu, and Sanjib Kumar Panda. Optimization of Distribution Network Incorporating Distributed Generators: An Integrated Approach. *IEEE TRANSACTIONS ON POWER SYSTEMS*, 28, Aug. 2013.

[65] J. A. Taylor and J. L. Mathieu. Index policies for demand response. *IEEE Transactions on Power System*, PP(99):1–9, 2013.

[66] R. J. Thomas, L. Tong, L. Jia, and O. E. Kosut. Some economic impacts of bad and malicious data. In *PSerc 2010 Workshop*, volume 1, Portland Maine, July 2010.

[67] F. F. Wu and W. E. Liu. Detection of topology errors by state estimation. *IEEE Transactions on Power Systems*, 4(1):176–183, Feb 1989.

[68] F. F. Wu, P. Varaiya, P. Spiller, and S. Oren. Folk theorems on transimission access: proofs and conterexamples. *Journal of Regulatory Economics*, 10, 1996.

[69] L. Xie, Y. Mo, and B. Sinopoli. False data injection attacks in electricity markets. In *Proc. IEEE 2010 SmartGridComm*, Gaithersburg, MD, USA., Oct. 2010.

[70] Y. Xu and L. Tong. On the value of storage at consumer locations. In *2014 IEEE PES General Meeting*, July 2014.

[71] P. Yang, G. Tang, and A. Nehorai. A game-theoretic approach for optimal time-of-use electricity pricing. *IEEE Transactions on Power Systems*, 2012.

[72] Y. Zhai, P. Tehrani, L. Li, J. Zhao, and Q. Zhao. Dynamic pricing under binary demand uncertainty: a multi-armed bandit with correlated arms. In *Proc. of the 45th IEEE Asilomar Conference on Signals, Systems, and Computers*, Nov. 2011.

[73] T. Zheng and E. Litvinov. Ex post pricing in the co-optimized energy and reserve market. *IEEE Transaction on Power System*, 21(4), 2006.