Demand Response Management for Profit Maximizing Energy Loads in Real-Time Electricity Market

Shuoyao Wang, Suzhi Bi, *Member, IEEE*, and Ying-Jun Angela Zhang, *Senior Member, IEEE*

***Abstract*—In this paper, we consider the profit-maximizing demand response of an energy load in the real-time electricity market. In a real-time electricity market, the market clearing price is determined by the random deviation of actual power supply and demand from the predicted values in the day-ahead market. An energy load, which requires a total amount of energy over a certain period of time, has the flexibility of shifting its energy usage in time, and therefore is in perfect position to exploit the volatile real-time market price through demand response. We show that the profit-maximizing demand response strategy can be obtained by solving a finite-horizon Markov decision process (MDP) problem, which requires extremely high computational complexity due to continuous state and action spaces. To tackle the high computational complexity, we propose a dual approximate approach that transforms the MDP problem into a linear programing problem by exploiting the threshold structure of the optimal solution. Then, a row-generation-based solution algorithm is proposed to solve the problem efficiently. We demonstrate through extensive simulations that the proposed method significantly reduces the computational complexity of the optimal MDP problem (linear versus exponential complexity), while incurring marginal performance loss. More interestingly, the proposed demand response strategy hits a triple win. It not only maximizes the profit of the energy load, but also alleviates the supply-demand imbalance in the power grid, and even reduces the bills of other market participants. On average, the proposed quadratic approximation and improved row generation algorithm increases the energy load's profit by 55.9% and saves the bills of other utilities by 80.2% comparing with the benchmark algorithms.**

***Index Terms*—Demand response, electricity market, energy load, power imbalance, MDP, column generation.**

Manuscript received September 6, 2017; revised January 3, 2018 and March 6, 2018; accepted April 10, 2018. Date of publication April 16, 2018; date of current version October 18, 2018. This work was supported in part by General Research Funding (Project number 14200315) from the Research Grants Council of Hong Kong and in part by the Theme-Based Research Scheme (Project number T23-407/13-N). The work of S. Bi was supported in part by the National Natural Science Foundation of China under Project 61501303, in part by the Foundation of Shenzhen City under Project JCYJ20160307153818306 and JCYJ20170818101824392, and in part by the Science and Technology Innovation Commission of Shenzhen under Project 827/000212. Paper no. TPWRS-01379-2017. *(Corresponding author: Suzhi Bi.)*

S. Wang and Y.-J. Angela Zhang are with the Information Engineering, The Chinese University of Hong Kong, 999077, Hong Kong (e-mail: w.shuoy@gmail.com; yjzhang@ie.cuhk.edu.hk).

S. Bi is with the Information Engineering, Shenzhen University, Shenzhen 518060, China (e-mail: bsz@szu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2018.2827401

NOMENCLATURE

Sets

$\ominus$	Set that defines the constraints set.
$\mathcal{A}$	Set of all possible actions.
$\mathcal{S}$	Set of all possible states.

Notation

λ_i	Day-ahead electricity price at hour i (\$/MWh).
λ_i^{RT}	Real-time electricity price at hour i (\$/MWh).
$g_i(P_i, w_i)$	Real-time electricity bill in hour i given P_i and w_i (\$).
$V_i(s)$	Value function of hour i given residual demand s (\$).
$\triangle d_i$	Net deviation at hour i (MWh).
$\mathcal{C}_i(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}, s_i, P_i)$	i-th constraint of RMP.
$\theta_i, \alpha_i, \beta_i$	Approximation parameters of the value function at hour i.
d_i	Day-ahead regulated volume of the grid at hour i (MWh).
E	Total demand in N hours (MWh).
$f(w)$	Probability density function of the coordinated imbalance w.
m_i	Total imbalance due to other customers at hour i (MWh).
p_i	Day-ahead regulated volume of the energy load at hour i (MWh).
s_i	Remaining demand of the energy load from hour i to N (MWh).
U	Actual power consumption upper bound (MWh).
w_i	Coordinated total imbalance at hour i.

Variable

P_i	Total energy flow to the energy load at hour i (MWh).
δ_i	Real-time customer imbalance (MWh).

I. INTRODUCTION

THE electricity markets in Europe and the U.S. have adopted a two-settlement structure, consisting of a day-ahead market and a real-time market, as illustrated in Fig. 1 [1]. The

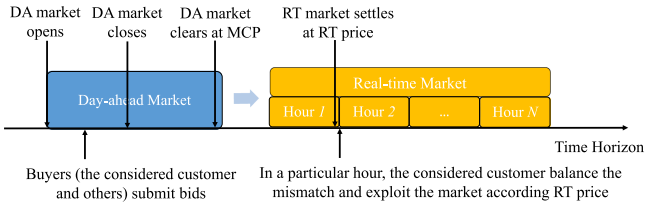


Fig. 1. Illustration of a two-settlement electricity market. DA: day-ahead market; RT: real-time market; MCP: market clearing price.

day-ahead market is a forward market, in which hourly electricity prices are calculated for each hour of the next operating day based on generation offers, demand bids, and scheduled bilateral transactions. Likewise, the real-time market is a spot market, in which real-time market prices are cleared based on deviations of the actual energy supply and demand from the day-ahead schedule. Due to the increasing penetration of renewable energy generation, real-time market prices are becoming more volatile and random.

The rising number of energy storage devices (e.g., batteries and electric vehicles) connected to the grid could lead to increasing participation of energy loads in the electricity market. Energy customers are the ones that require a total amount of energy over a given period of time, regardless of how much power is consumed at particular time instants. In practice, an energy load can be an electric vehicle (EV) aggregator, a grid-connected energy storage, etc. Being able to freely shift the power usage over time, energy loads are in the perfect position to exploit the volatile real-time electricity price through demand response (DR).

DR has been extensively studied as a mechanism for electricity customers to adjust their power consumption, so that the power demand is better matched with the supply. The benefits of DR include increased system stability [2], reduced capital expenditures for peak load demand [3], reduced average power generation costs [4], lowered electricity bills for end users [5], etc. Methods of engaging users in DR efforts include offering time-of-use pricing, critical peak pricing, critical peak rebate, real-time pricing, etc. [6]. For example, [5] proposed a real-time pricing scheme that maximizes the financial benefits of both the utility and users, where the optimal price is calculated to incentivise users to adjust their power consumption. The loads in DR programs are typically classified into time-shiftable loads, i.e., fixed-size deferrable loads with deadlines, (e.g., data-center workloads [3], washer [7]) and elastic loads, i.e., variable-size non-deferrable loads, (e.g., data-center cooling [3] , lighting [8]). With the development of advanced battery technology, increasingly energy loads (e.g., energy storage systems [2], EV [9]) start to participate in demand response programs. Energy loads, which requires a total amount of energy over a certain period of time, has the flexibility of reshaping its energy consumption profile within the time period.

DR has recently been applied to local electricity markets [10]–[20]. There are two main types of local electricity markets: day-ahead and real-time markets. For day-ahead markets, various algorithms have been proposed to investigate the optimal bidding strategy through mixed-integer linear programming [10], game theory [11], bi-level stochastic programming [12], etc. Besides, [13] studied the impact of price-based DR programs on market prices through simulating a day-ahead electricity market with trading agents. Sedzro *et al.* [14] investigated how a cooperative group of customers pool together their deferral loads and participate in the day-ahead market via an aggregator. Alternatively, the real-time market has also been taken into consideration [15]–[20]. For instance, [15], [16] proposed price arbitrage strategies for energy storage devices in a real-time electricity market. Muratori and Rizzoni [17] investigated an optimal residual appliance schedule problem using dynamic programming that minimizes a cost function under time-of-use electricity price. The above work [15]–[17] assumed that the real-time market prices are given as exogenous variables and the DR participants are price-takers. In practice, however, a large-scale energy load can affect the market price. Vespermann *et al.* [18]–[20] studied the offering and operation strategies of large energy storage units that can affect the market price. However, [18], [19] are based on a non-realistic assumption that the non-causal information of future supply-demand profiles is perfectly known at the beginning of system time. In contrast, Kohansal and Mohsenian-Rad [20] investigated the hourly-independent economic bidding and DR problem of times-shiftable loads through stochastic programming.

To complement most of the previous work that does not consider the sequential procurement and uncertain future information in real-time markets, we investigate the optimal multi-stage DR strategy that maximizes the profit of a large-scale price-maker energy load in the real-time electricity market. In particular, the energy load reshapes its power consumption profile over time in response to the real-time deviation of power supply and demand with respect to the predicted values in the day-ahead market. Due to the DR activity, the real-time power supply-demand imbalance is reshaped, thus affecting the electricity price in the real-time market. Here, the DR optimization problem is naturally formulated as a stochastic Markov decision process (MDP) problem due to the following two reasons. First, the constraint on the total energy demand couples the decisions in different hours. Secondly, the energy load does not have the non-causal knowledge of the realizations of future supply-demand imbalance. Instead, only statistical distributions can be obtained from historical data. This is in contrast to the assumptions in [18], [19]. As such, the energy load needs to make sequential decisions, where each decision is made based on the causal information available until that time instant. Nonetheless, the MDP problem formulated is challenging in the sense that the state and action spaces are continuous, because the total energy demand and the energy transaction amount at each time are continuous real numbers. The problem thus suffers the curse of dimensionality and is in general hard to solve. Our main contributions are summarized as follows.

- To the best of our knowledge, this is the first paper that considers the multi-stage demand response of a price-maker energy load in the real-time electricity markets, where the energy load reshapes its energy consumption profile

sequentially in response to the power supply-demand imbalance of each hour. In particular, the problem is naturally a stochastic MDP with continuous state and action spaces.
- Through rigorous analysis, we find the threshold structure of the optimal DR solution. Based on this structure, we can solve the MDP problem through backward induction regardless of the continuous state and action spaces.
- To further reduce the complexity of backward induction, we propose a quadratic-approximation formulation, which dramatically reduces the scale of the dual problem of the above MDP problem. The number of variables in each stage is reduced from infinite to three.
- To efficiently solve the reduced dual problem, we derive an improved row generation [21] based algorithm, referred to as QARG. The proposed algorithm significantly reduces the complexity of solving the MDP problem from exponential to linear with respect to the length of the time horizon. We use real-world data to evaluate the proposed QARG strategy and show that QARG hits a triple win. On average, QARG increases the energy load's profit by 55.9% and saves the bills of other utilities by 80.2% comparing with a benchmark greedy algorithm. Moreover, QARG also reduces the mean squared imbalance of the grid, where the reduction is 291.5% more than the reduction of the greedy algorithm.

The rest of this paper is organized as follows. In Section II, we introduce the system model. In Section III, we derive a threshold structure optimal solution to solve the MDP problem. We propose a quadratic approximation formulation and an efficient row generation based algorithm to obtain a near-optimal operation strategy in Section IV. Numerical results and conclusions are presented in Sections V and VI.

II. System Models

A. Market Model

We consider a two-settlement coupled electricity market, consisting of a day-ahead market and a real-time market (Fig. 1). The day-ahead market is a forward market, in which electricity prices are calculated for each hour of the next operating day based on the submitted generation offers, demand bids, virtual supply offers, virtual demand bids and bilateral transaction schedules. Suppose that we are interested in a time period of N hours within an operation day. Denote by d_i (MWh) the load demand cleared[1] in the day-ahead market at hour i, and denote by λ_i ($/MWh) the corresponding clearing price.

On the other hand, the real-time market is a balancing market, in which the electricity prices are calculated based on the actual system operations and the day-ahead regulations. The actual energy supply and demand may deviate from the predicted values in the day-ahead market. Suppose that the net deviation at hour i is $\triangle d_i$. In particular, $\triangle d_i > 0$ implies that the actual demand exceeds the actual energy supply and the grid issues a higher electricity price that encourages the generators to produce more energy and penalizes the excessive loads, and vice versa, in the hope to reduce the imbalance.

The real-time electricity price is determined by solving a congestion-aware optimization problem to obtain the locational marginal prices (LMPs) based on the estimation of actual power generation and load demand status, and also other network physical characteristics [23]. In particular, given the day-ahead clearing price λ_i, the real-time clearing price could be interpreted as a function of the fraction $\frac{\triangle d_i}{d_i}$ [20], [22]. To capture the impact of demand response to the real-time electricity price, we consider a general monotonic non-decreasing $h_i(\cdot)$ to model the real-time price as a function of the real-time load imbalance:

$$\lambda_i^{\mathrm{RT}} = h_i\left(\frac{\triangle d_i}{d_i}\right). \tag{1}$$

In practice, the general price function covers a wide variety of electricity pricing models influenced by either the physical components or the regulating policies of a power grid.

B. Energy Load

We consider an energy load, e.g., a group of EV via an aggregator or a group of energy storage systems, that requires a total of E MWh energy within a certain period time. To satisfy the demand, the energy load buys from the day-ahead market p_i (MWh) electricity at hour i. In real-time operation, the energy load observes the power deviation from day-ahead dispatch due to other customers, and adjusts its own power consumption by selling or buying in the real-time market. Let δ_i (MWh) denote the amount of power the energy load buys ($\delta_i > 0$) or sells ($\delta_i < 0$) in the real-time market at hour i. Then, the actual power consumption of the energy load at hour i is $P_i = p_i + \delta_i$. We restrict P_i to be non-negative, implying that the energy load cannot feed power back to the grid. This is in accordance to the regulation of existing markets, e.g., PJM and Nordic markets. Moreover, P_i is upper bounded by a constant U (MWh) due to the physical constraints of the distribution network and the storage device. To fulfill the energy demand within N hours, we use $V_{N+1}(s)$ to denote the dissatisfaction imposed to the energy load at the end of the N-th hour for not being able to satisfy all the demand, where $V_{N+1}(0) = 0$ and $V_{N+1}(s)$ is a non-decreasing convex function to indicate that the energy load fulfills the demand as much as possible. Recall that the real-time market price is closely related to the power imbalance from the day-ahead schedule, i.e., (1). Suppose that the imbalance due to other suppliers (e.g., uncertain renewable energy) and customers (e.g., DR of other demand) at hour i is denoted by m_i. Then, we have $\triangle d_i = m_i + \delta_i$ and $\lambda_i^{\mathrm{RT}} = h_i(\frac{m_i+\delta_i}{d_i})$. As a result, the total real-time cost for the energy load is

$$\sum_{i=1}^{N} h_i\left(\frac{m_i+\delta_i}{d_i}\right)\delta_i. \tag{2}$$

Intuitively, with DR in the real-time market, the cost to the energy load must be lower than in the case without DR, i.e., be negative. The additive inverse of the real-time negative cost can be viewed as a profit the energy load derives from the real-time market manipulation.

[1]The cleared generation is equal to the cleared demand in the day-ahead market.

C. Problem Formulation

At the beginning of each hour i, the optimal δ_i is determined based on the residual demand s and the observation of m_i. For notation simplicity, we denote $w_i = \frac{m_i + d_i}{m_i}$ and the real-time electricity bill in hour i as $g_i(P_i, w_i) = h_i(w_i)(P_i - p_i)$. The probability density function (PDF) of w_i is denoted as $f(w)$. When making the decision, the energy load has no prior knowledge of the realizations of m_k for $k = i, i+1, ..., N$. As such, the problem is formulated as the following stochastic MDP problem [24], where the residue energy demand s is regarded as the system state and the state transition from i to $i+1$ is determined by the decision δ_i as well the exogenous variables p_i.

$$s \leftarrow s - \delta_i - p_i. \tag{3}$$

The state space $\mathcal{S}$ and action space $\mathcal{A}$ are $[0, E]$ and $[0, U]$, respectively. At stage k, the objective function to solve is,

$$\begin{aligned} &V_i(s, w_i) = \\ &\min_{\delta_i} \quad g_i(p_i + \delta_i, w_i) + E_{w_{i+1}}\left[V_{i+1}(s - \delta_i - p_i, w_{i+1})\right] \\ &\text{s.t.} \quad -p_i \le \delta_i \le U - p_i, \end{aligned} \tag{4}$$

for $i = 1, ..., N$, where $V_{N+1}(s, w_i) = V_{N+1}(s)$.

We assume that m_i's are independent random variables (the assumption is verified in Simulation) whose distributions can be estimated from the historical data. For simplicity of illustration, we substitute δ_i by $P_i - p_i$. Accordingly, for $i = 1, \ldots, N$, we can obtain the value functions of the above MDP problem from (4) by taking expectation over w_i,

$$\begin{aligned} V_i(s) = E_{w_i}\left[\min_{P_i} \quad g_i(P_i, w_i) + V_{i+1}(s - P_i)\right] \\ \text{s.t.} \quad 0 \le P_i \le U. \end{aligned} \tag{5}$$

With the above formulation, we are interested in the optimal strategies δ_i's that yield the minimum expected total bill $V_1(E)$. In the following sections, we introduce the methods to find the optimal decision rule and the value of $V_1(E)$.

III. Optimal Station Operation

Finite horizon MDP problems can typically be solved by backward induction, where the complexity is $O(|\mathcal{S}|^N |\mathcal{A}|^N)$ [24]. However, in this paper, the state space $\mathcal{S} = [0, E]$ and the action space $\mathcal{A} = [0, U]$ are continuous intervals and discretizing the spaces may lead to undesirable drawbacks. First, the fine granularity of discretization causes high dimensions of state and action spaces, which in turn leads to prohibitively high computational complexity [2, Remark 1]. Secondly, discretization errors may cause issues with the convergence or stability of the MDP algorithms [25]. As a consequence, there always exists a conflict between control accuracy and required training time. To tackle these problems, in this section, we show that the optimal cost-to-go function is convex non-decreasing in s and the optimal solution has a threshold structure. This enables us to apply backward induction without state and action discretization.

Lemma 1: For every $s \in \mathcal{S}$ and $i = 1, \ldots, N$, $V_i(s)$ is a convex non-decreasing function in s.

Proof: a) Convexity: we prove the the convexity of $V_i(s)$ by induction. i) The terminal cost $V_{N+1}(s)$, is a convex function by definition. ii) Suppose that $V_{i+1}(s)$ is convex for any $i \le N$. The joint convexity of $g_i(P_i, w_i)$ in (P_i, w_i) is straightforward from the definition. This implies the joint convexity of $g_i(P_i, w_i) + E_{w_{i+1}}[V_{i+1}(s - P_i, w_{i+1})]$ in (P_i, w_i). Then, $V_i(s, w_i)$ is convex in s. As a weighted sum of $V_i(s, w_i)$, $V_i(s)$ is also convex in s. i) and ii) together conclude the proof.

b) Non-decreasing: we prove the non-decreasing monotonicity of $V_i(s)$ by induction. i) The terminal cost $V_{N+1}(s)$ is a non-decreasing function by definition. ii) Suppose that $V_{i+1}(s)$ is non-decreasing for any $i \le N$. For each s and w_i, we denote the optimal decision as P_{s,w_i}. Then, for a small $\phi > 0$, we have $V_{i+1}(s + \phi) + g_i(P_{s,w_i}, w_i) \ge V_{i+1}(s) + g_i(P_{s,w_i}, w_i)$. According to the optimality of $P_{s+\phi,w_i}$, we have $V_{i+1}(s + \phi) + g_i(P_{s+\phi,w_i}, w_i) \ge V_{i+1}(s + \phi) + g_i(P_{s,w_i}, w_i)$. Substitute the inequalities into (3), for any small $\phi > 0$, we have

$$\begin{aligned} &V_i(s + \phi) \\ &= \int \left[V_{i+1}(s + \phi - P_{s+\phi,w}) + g_i(P_{s+\phi,w}, w)\right] f(w) dw \\ &\ge \int \left[V_{i+1}(s - P_{s,w}) + g_i(P_{s,w}, w)\right] f(w) dw = V_i(s). \end{aligned} \tag{6}$$

The statements in i) and ii) together conclude the proof. ■

We denote the right and left derivatives of $V_i(s)$ as $h_i^+(s) \triangleq \lim_{\phi \to 0} \frac{V_i(s+\phi) - V_i(s)}{\phi}, \forall \phi \ge 0$ and $h_i^-(s) \triangleq \lim_{\phi \to 0} \frac{V_i(s+\phi) - V_i(s)}{\phi}, \forall \phi \le 0$, respectively. In the following Theorem 1, we derive a threshold structure of the optimal policy from Lemma 1. The threshold is to compare the derivative of $g_i(s, w_i)$ with that of $V_i(s, w_i)$ for given a w_i.

Theorem 1: The threshold structure of the optimal policy for hour i is,

$$P_i^* = \begin{cases} U, \text{ if } h_{i+1}^-(s - U) \ge \dfrac{\partial g_i(U, w_i)}{\partial P} \\ 0, \text{ if } h_{i+1}^+(s) \le \dfrac{\partial g_i(0, w_i)}{\partial P} \\ x, \text{ otherwise,} \\ \text{where } x = \sup\left\{x : h_{i+1}^+(s - x) \ge \dfrac{\partial g_i(x, w_i)}{\partial P}\right\}. \end{cases} \tag{7}$$

Due to page limit, we omit the proof here. The main idea of the proof is as follows. Lemma 1, together with the fact that $g_i(P_i, \omega_i)$ is a quadratic function, implies that Problem (5) is a convex problem with linear constraints. Consequently, the Karush-Kuhn-Tucker (KKT) conditions are the necessary and sufficient conditions for optimality. By manipulating the KKT conditions of (5), we have Theorem 1.

Intuitively, P_i^* is the value that achieves the balance between the expected future bill and the current bill. In particular, if the marginal expected future bill is no-larger than the marginal current bill $P_i^* = 0$ and if the marginal expected future bill is no-less than the marginal current bill $P_i^* = U$. Furthermore,

Algorithm 1: Backward Induction to Find Optimal $\{P_i^*\}$.

Initialization: $V_{n+1}(s)$, $f(w)$, $h_{n+1}^-(s)$, $h_{n+1}^+(s)$, $\{d_i\}, \{\lambda_i\}, \{p_i\}$

1: **for** $i = N$ down to 1 **do**
2: Compute P_i^* using (7)
3: Compute $V_i(s, w_i)$ using (8)
4: Take the integral of $V_i(s, w_i)$ over w_i and then take the derivative of $V_i(s)$ to get $h_i^+(s)$ and $h_i^-(s)$
5: **end for**

Output: $\{P_i^*\}$

substituting (7) into (4), we have,

$$V_i(s, w_i) = \begin{cases} g_i(U, w_i) + V_{i+1}(s-U), \text{if } h_{i+1}^-(s-U) \geq \dfrac{\partial g_i(U, w_i)}{\partial P} \\ g_i(0, w_i) + V_{i+1}(s), \text{if } h_{i+1}^+(s) \leq \dfrac{\partial g_i(0, w_i)}{\partial P} \\ g_i(x, w_i) + V_{i+1}(s-x), \text{otherwise,} \\ \text{where } x = \sup\left\{ x : h_{i+1}^+(s-x) \geq \dfrac{\partial g_i(x, w_i)}{\partial P} \right\}. \end{cases} \tag{8}$$

Thanks to (7) and (8), we can find the optimal $\{P_i^*\}$ through Algorithm 1.

According to (8), $V_i(s, w_i)$ is a piecewise-defined function. The number of sub-functions increases by three times when we process from $V_{i+1}(s)$ to $V_i(s, w_i)$. Therefore, although the threshold structure approach provides an optimal solution without the need of space discretization, the computational complexity of $V_i(s)$ still increases with the number of stages exponentially. This renders backward induction unscalable with the length of the time horizon. To address this issue, we propose a reduced-complexity suboptimal algorithm based on a quadratic approximation of $V_i(s)$ in Section IV.

IV. Quadratic Approximation and Improved Row Generation Algorithm

In this section, we propose an approximate algorithm, referred to as QARG, to reduce the complexity of solving the MDP problem. In particular, we apply a quadratic approximation on the dual problem of (5). Then, the problem is solved by an improved row-generation method.

A. Dual Problem and Quadratic Approximation

According to [24], the minimum expected total bill $V_1(E)$ in (5) can be obtained by solving the following dual problem,

$$\begin{aligned} \max_{\{P_i\}} \quad & V_1(E) \\ \text{s.t.} \quad & V_i(s) - \int \left[g_i(P_i, w) + V_{i+1}(s - P_i) \right] f(w) dw \\ & \leq 0, \forall s \in \mathcal{S}, \forall i = 1, \ldots, N \\ & V_i(s) \in \mathbb{R}, \forall s \in \mathcal{S}, \forall i = 1, \ldots, N \\ & 0 \leq P_i \leq U. \end{aligned} \tag{9}$$

In general, the dual problem does not reduce the complexity of MDP problems. One effective method to reduce the complexity is to approximate $V_i(s)$ by a polynomial function [25]. As $g_i(P, w)$ is a quadratic function of P and $V_i(s)$ is convex non-decreasing in s, it is reasonable to approximate $V_i(s)$ by a quadratic function

$$V_i(s) \approx \theta_i + \alpha_i s + \beta_i s^2, \forall i = 1, \ldots, N. \tag{10}$$

If θ_i, α_i, and β_i are known parameters, we can simplify the decision rules in Theorem 1. Taking the derivative of (10), we have $h_{i+1}^-(s) = h_{i+1}^+(s) = \alpha_i + 2\beta_i s$. Substituting it into Theorem 1, we have the following Proposition 1.

Proposition 1: If the cost-to-go function of stage i is $\theta_i + \alpha_i s + \beta_i s^2$, the optimal decision at stage i is

$$P_i^* = \left[\frac{-\lambda_i w_i + \frac{2\lambda_i p_i}{d_i} + (\alpha_{i+1} + 2\beta_{i+1}) s}{2\beta_{i+1} + \frac{2\lambda_i}{d_i}} \right]_0^U \tag{11}$$

Proposition 1 specifies the fixed decision rule without any recursion when the parameters $\{\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}\}$ are given, where $\boldsymbol{\theta} \triangleq (\theta_1, \ldots, \theta_N)$, $\boldsymbol{\beta} \triangleq (\beta_1, \ldots, \beta_N)$, and $\boldsymbol{\alpha} \triangleq (\alpha_1, \ldots, \alpha_N)$. In the following, we propose an efficient algorithm to calculate these parameters. For notation simplicity, we define

$$\begin{aligned} \mathcal{C}_i(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}, s, P) \triangleq & \theta_i - \theta_{i+1} + (\alpha_i - \alpha_{i+1}) s + (\beta_i - \beta_{i+1}) s^2 \\ & + \int \left[\alpha_{i+1} P - g_i(P, w) + \beta_{i+1} \left(2Ps - P^2 \right) \right] f(w) dw, \end{aligned} \tag{12}$$

for $i = 1, \ldots, N-1$ and

$$\begin{aligned} \mathcal{C}_N(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}, s, P) \triangleq & \theta_N - \theta_{N+1} \\ & + (\alpha_N - \alpha_{N+1}) s + (\beta_N - \beta_{N+1}) s^2 \\ & + \int \left[\alpha_{N+1} P - g_N(P, w) + \beta_{N+1} \left(2Ps - P^2 \right) \right] f(w) dw. \end{aligned} \tag{13}$$

Substituting (10) into (9) yields an approximated linear programming problem as follows.

$$\begin{aligned} \max_{\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}} \quad & \theta_1 + \alpha_1 E + \beta_1 E^2 \qquad \text{[ALP]} \\ \text{s.t.} \quad & \mathcal{C}_i(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}, s_i, P_i) \leq 0, \forall s \in \mathcal{S}, \forall P_i \in [0, U], \\ & i = 1, \ldots, N. \end{aligned} \tag{14}$$

The approximated linear programming [ALP] reduces the number of variables in (8) from infinite to $3N$ with the same number

of constraints. Nonetheless, the number of constraints remains infinite due the continuous state space. Row generation is a powerful tool for solving large-scale linear programming problems. However, the standard cutting problem of (14) is still a large-scale linear programming and thus direct application of standard row generation to (14) may suffer a large number of iterations and a long convergence time [26]. In the following, we develop an improved row generation based solution approach to handle the challenge [27].

B. Improved Row Generation Based Algorithm

The framework of the proposed algorithm is to solve [ALP] by iteratively solving two sub-problems, i.e., Restricted Master Problem [RMP] and Cutting Problem [CP]. [RMP] is formed by replacing the constraints in [ALP] by a small but meaningful subset of the constraints. The optimal values of [RMP] are used to determine whether there exist feasible (s_i, P_i)'s that can cut out some infeasible solutions of [ALP]. This procedure is formulated as [CP]. We name (s_i, P_i) as a "row", because (s_i, P_i) adds constraints to [RMP] to cut out infeasible solutions. If the potential cutting row(s) exist, we add the corresponding constraints to the subset and re-solve the new [RMP]. This process repeats until no new rows can be found [27]. The details of [RMP] and [CP] are as follows.

1) Restricted Master Problem: Without loss of generality, we start with the initial set $\ominus = \{(s_i, P_i) | P_i = p_i, s_i = E - \sum_{j=1}^{i-1} p_j, \forall i = 1, \ldots, N\}$ and solve (13) with a set of constraints defined by $(s_i, P_i) \in \ominus$:

$$\begin{aligned} \max_{\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}} \quad & \theta_1 + \alpha_1 E + \beta_1 E^2 && \text{[RMP]} \\ \text{s.t.} \quad & \mathcal{C}_i(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}, s_i, P_i) \le 0, \forall (s_i, P_i) \in \ominus, \\ & i = 1, \ldots, N. && (15) \end{aligned}$$

The optimal value of [RMP] is an upper bound of the one of [ALP], as the constraints of [RMP] are a subset of the constraints in [ALP]. Notice that the optimal solution may not be unique. In this case, we consider a non-decreasing-β optimal solution $(\boldsymbol{\theta}^0, \boldsymbol{\alpha}^0, \boldsymbol{\beta}^0)$, i.e., $\beta_i^0 \le \beta_{i+1}^0 \; \forall \; i = 1, \ldots, N$. Besides, Lemma 2 guarantees the existence of the non-decreasing-β solution, which ensures the concavity of [CP].

Lemma 2: There is always an optimal solution $(\boldsymbol{\theta}^0, \boldsymbol{\alpha}^0, \boldsymbol{\beta}^0)$ that the entries of $\boldsymbol{\beta}^0$ are always in non-decreasing order, i.e., $\beta_i^0 \le \beta_{i+1}^0 \; \forall i = 1, \ldots, N$.

Proof: We assume that there exists an optimal solution $(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ where $\beta_j < \beta_{j-1}$. We denote $\boldsymbol{\beta}' = (\beta_1, \ldots, \beta_{j-2}, \beta_{j-1}, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_N)$. Substituting $\boldsymbol{\beta}'$ into [ALP], we have that $(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}')$ also satisfies all the constraints and achieves the same $\theta_1 + \alpha_1 E + \beta_1' E^2$. Repeating the above process, for any set of optimal solutions, we can always find a non-decreasing $\boldsymbol{\beta}^0 = (\beta_1^0, \beta_2^0, \ldots, \beta_N^0)$ where $(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}^0)$ is also an optimal solution. ■

2) Cutting Problem: Our next task is to find a (s_i, P_i) such that $\ominus \cup (s_i, P_i)$ results in a tighter upper bound of the [ALP] than the one with $\ominus$. Specifically, the typical cutting problem finds a row with the largest $\mathcal{C}_1$ and non-positive $\mathcal{C}_2, \ldots, \mathcal{C}_N$ [27]. In fact, it is not necessary to select the largest $\mathcal{C}_1$. Given $(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta})$, any row with positive $\mathcal{C}_i$, i.e., violating the constraint in (13), is a candidate to be a row [28]. In particular, given the optimal solution $(\boldsymbol{\theta}^0, \boldsymbol{\alpha}^0, \boldsymbol{\beta}^0)$ of (RMP), we find the row (s_i, P_i) that maximize the $\mathcal{C}_i$,

$$\begin{aligned} \max_{s, P} \quad & \mathcal{C}_i(\boldsymbol{\theta}^0, \boldsymbol{\alpha}^0, \boldsymbol{\beta}^0, s, P) && [\text{CP}]_i \\ \text{s.t.} \quad & s \in \mathcal{S}, \\ & 0 \le P \le U, && (16) \end{aligned}$$

for $i = 1, \ldots, N$. The coefficient of quadratic terms in the objective function of (16), i.e., $\beta_i - \beta_{i+1}$ and $-\beta_{i+1}$ are both negative according to Lemma 2. By calculating the determinant of the Hessian matrix, we can show that $[\text{CP}]_1$, ..., $[\text{CP}]_N$ are concave maximization problems. The KKT conditions of $[\text{CP}]_i$ are

$$\begin{aligned} & \left(\alpha_i^0 - \alpha_{i+1}^0\right) + 2\left(\beta_i^0 - \beta_{i+1}^0\right) s + 2\beta_{i+1}^0 P + \mu_1 - \mu_2 = 0, \\ & \alpha_{i+1}^0 - \lambda_i + \frac{2\lambda_i P}{d_i} - \frac{2\lambda_i p_i}{d_i} - \beta_{i+1}^0 (2P - 2s) + \omega_1 - \omega_2 = 0, \\ & \mu_1 s = 0, \; -\mu_2 (s - s_{\max}) = 0, \; \omega_1 P = 0, \; -\omega_2 (P - U) = 0, \end{aligned} \tag{17}$$

where μ_1, μ_2, ω_1, and ω_2 are the Lagrange multipliers. The row (s, P) is solved by the boundary conditions of (s, P). If the maximum value of $[\text{CP}]_i$ is positive, we update $\ominus = \ominus \cup (s, P)$. If $|\ominus|$ increases after checking $[\text{CP}]_1$, ..., $[\text{CP}]_N$, we return a new RMP. Otherwise, at this point, the optimal solution to [RMP] satisfy all the constraints of [ALP] and thus is the optimal solution to [ALP]. We terminate the algorithm.

On the one hand, because the objective function and constraints are only related to $(\boldsymbol{\theta}^0, \boldsymbol{\alpha}^0, \boldsymbol{\beta}^0)$ and $\ominus$, the update of $\mathcal{C}_i$ are independent. Therefore, we can solve $[\text{CP}]_1$, ..., $[\text{CP}]_N$ parallelly, which greatly reduces the computational time. On the other hand, each update reduces the objective value in (15), and thus convergence is guaranteed. The algorithm described above is summarized in Algorithm 2.

V. Simulation Results

In this section, we use real-world data to evaluate the performance of QARG. Without loss of generality, we consider a 10-hour time period (from 9:00 to 18:00) in all simulations. All the simulation results are the average performance of 90 days. The upper bound of the actual demand in each hour U is set as 50 MWh throughout the simulations.

A. Experimental Setup

We base our simulations on the historical hourly data, i.e., the regulated bidding volumes of day-ahead market d_i (Fig. 2(a)), the day-ahead electricity prices λ_i (Fig. 2(b)), and the real-time imbalances m_i (Fig. 2(c)) of the Finland Grid in Nordic electricity market [29]. We use the scaled regulated bidding volume of Finland-to-Estonia (Fig. 2(d)) in Nordic electricity market to model the energy load's regulated bidding volumes of each hour p_i, where the average total regulated bidding volume is set equal to the total demands in different demand scenarios

Algorithm 2: Improved Row Generation Based Algorithm.

Initialization: $\ominus \leftarrow \{(s_i, P_i)|P = p_i, s_i = E - \sum_{j=1}^{i-1} p_j, \forall i = 1, \ldots, N\}$, $F \leftarrow \mathbf{true}$

do
  $Z \leftarrow |\ominus|$.
  $(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \leftarrow$ the optimal solution of [RMP] given $\ominus$
  Formulate $[\text{CP}]_1$, ..., $[\text{CP}]_N$ with the optimal $(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta})$
  for $i = 1$ up to N **do**
    $(s_i^*, P_i^*) \leftarrow$ the optimal solution of $[CP]_i$
    $v_i^* \leftarrow$ the optimal value of $[CP]_i$
    if $v_i^* > 0$ **then**
      $\ominus = \ominus \cup (s_i^*, P_i^*)$
    end if
  end for
  if $|\ominus| \neq Z$ **then**
    $F \leftarrow$ **false**;
  end if
while F

Output: $(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta})$

(e.g., Fig. 5). The data set spans the first 3 months of 2017. The cumulative density function of m_i during this time period is plotted in Fig. 2(e). For simplicity, in our simulation, we use a linear function, i.e., $\lambda_i^{\text{RT}} = \lambda_i(\frac{m_i+\delta_i}{d_i})$, as the real-time price function [30]. For performance comparison, we consider the optimal backward induction solution (BI) [13], [20], the Greedy policy [17], and heuristic-greedy policy (HG) [14] as benchmarks.

- Greedy policy: In each stage, the energy load makes the buying/selling decision that minimizes the energy bill by ignoring the possible DR operation in the rest of time, i.e., setting $\delta_j = 0$, for $j = i+1, \ldots, N$. In particular, the energy load make the decision $\delta_i^{\text{Greedy}}(s) = \arg\min_\delta \lambda_i(w_i + \frac{1}{d_i}\delta)\delta + V_{N+1}(s - \sum_{j=i+1}^{N} p_j - \delta)$.
- HG: In each stage, the energy load makes the decision $\delta_i^{\text{HG}}(s) = \arg\min_\delta \lambda_i(w_i + \frac{1}{d_i}\delta)\delta + \epsilon V_{N+1}(s - \sum_{j=i+1}^{N} p_j - \delta)$, where $\epsilon = \epsilon_1$ when $w_i > 1$, $\epsilon = \epsilon_2$ when $w_i < 1$, and $\epsilon = 1$ otherwise. The selection of the value of the predetermined weights ϵ_1 and ϵ_2 are numerically studied and set as 1.02 and 0.96 in our experiment, respectively.

B. Independence Validation

An underlying assumption of our analysis is that m_i is i.i.d. for different i and orthogonal to λ_i. From Fig. 3(c), m_i is a first-order stationary process with zero mean. To validate the assumption of independence, we plot the auto-correlations and cross-correlations of the variables in Figs. 3(a) and (b), respectively. As we can see from Fig. 3(a), the auto-correlations of the variables reach the peak when the time lag is 0 and are close to zero at non-zero time lags, implying that m_i's are approximately mutually independent. Likewise, Fig. 3(b) shows that the cross-correlations of the variables are all close to zero, implying that m are approximately orthogonal to λ.

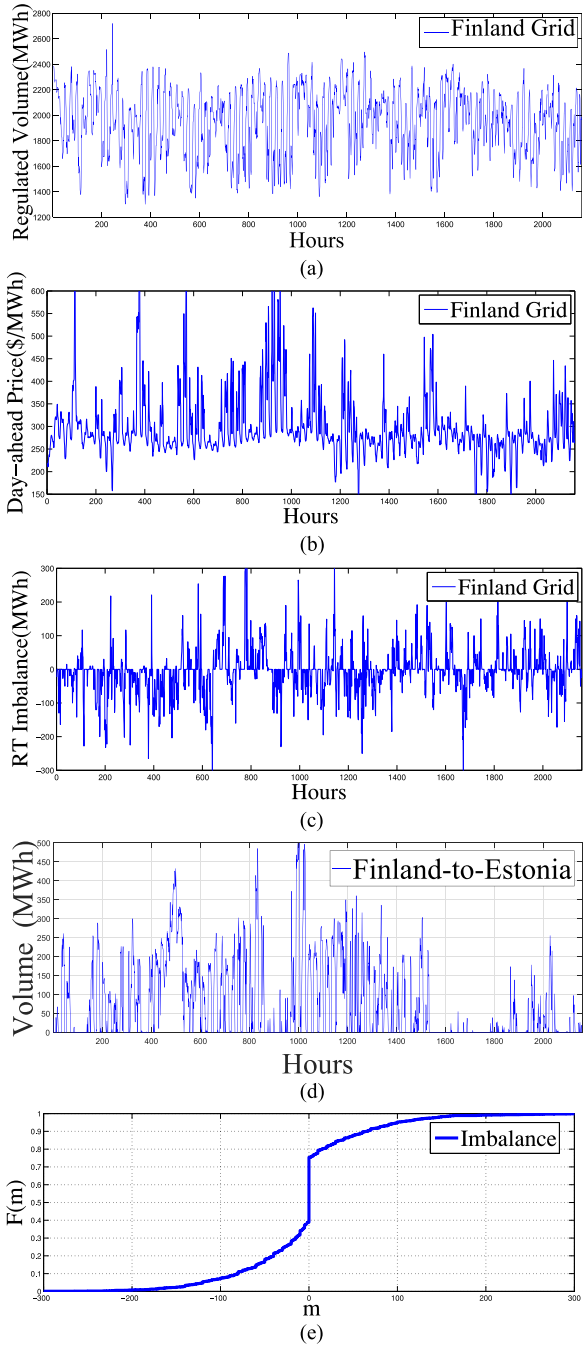


Fig. 2. Data for conducting simulations. (a) The bidding volumes of day-ahead market of Finland Grid. (b) The day-ahead hourly prices of Finland Grid. (c) The hourly load imbalance of Finland Grid. (d) The scaled real data regulated volume of Finland to Estonia. (e) Empirical cumulative distribution function of m.

C. Profit Performance

In this subsection, we investigate how the demand response strategy and profit of energy load versus the total demand E.

In Fig. 4, we show the optimized demand response using data from 2017/02/01 to 2017/02/03 as an example. Overall, we observe that when the real-time imbalance is negative (Fig. 4(a)), the energy load buys extensive power from the grid, i.e., $\delta > 0$,

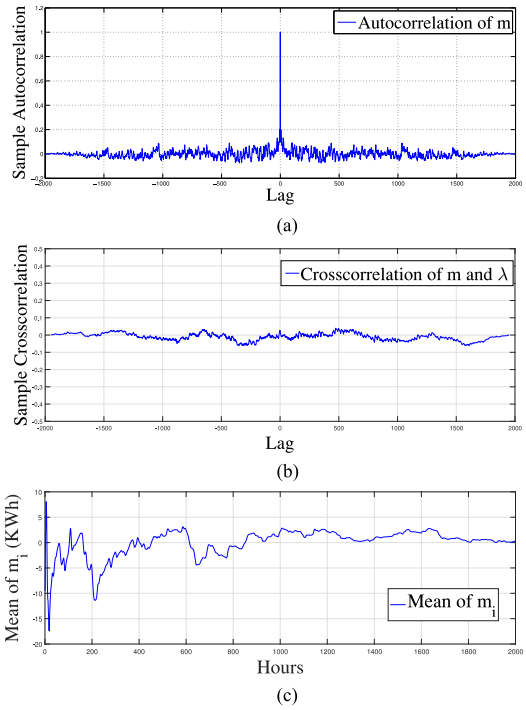


Fig. 3. Independence validation. (a) Auto-correlations of m. (b) Cross-correlations of m and λ. (c) Mean of m.

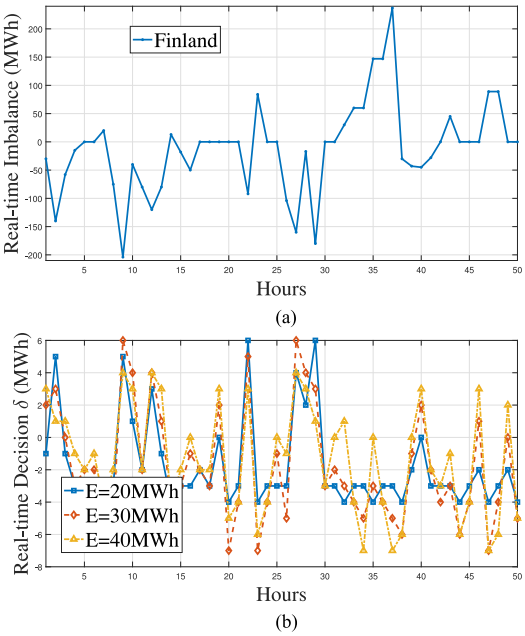


Fig. 4. Optimized demand response decision examples. (a) Real-time imbalance from 2017/02/01 to 2017/02/03. (b) Real-time decision from 2017/02/01 to 2017/02/03.

and vice versa. When the total demand and day-ahead regulated volume is large, the real-time decision is more aggressive in terms of the amount of real-time operation is much larger than the amount when E is small. This is because the selling operation is upper bounded by the day-ahead trading volume. In brief, the inverse correlation between the optimal demand strategy and the real-time imbalance indicates that QARG reduces the imbalance of the grid.

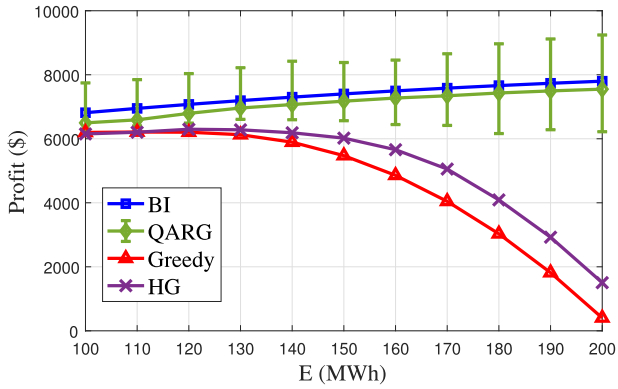


Fig. 5. Average total bill evaluation versus E.

In Fig. 5, we compare the average profit per day achieved by the three schemes with the same day-ahead bid realization, i.e., same day-ahead bidding volumes. Besides, we also plot the 95% confidence interval of QARG. In particular, we vary the total demand E from 100 MWh to 200 MWh and evaluate the each day's profit. For each value of E, the day-ahead bid realization is scaled so that the mean of accumulated bid realizations in each day is equal to E. The figure shows that QARG performs closely to the optimal BI approach and the performance gap does not increase with E. Overall, QARG outperforms the average greedy method profit performance with confidence level 95%. On average, the gap between QARG and BI is only 1.36% to 4.77% of the average profit achieved by BI. QARG achieves 55.9% and 50.1% higher profit than the Greedy method and HG, respectively. Furthermore, the percentile loss decreases as the load increases. In contrast, the Greedy method and HG perform poorly with a widening performance gap over the increase of E. This is because, when E is large, the Greedy method and HG tend to buy excessive energy at over-load period to eagerly satisfy E, which leads to a very high bill due to high electricity price. The proposed QARG method and BI can anticipate the future time-varying electricity price, thus can manage the electricity bill through buying energy only at off-peak hours and selling energy at overload period.

D. Impact to the Grid

The large grid supply-demand imbalance can result in large generation cost, large power loss, system security issues, etc. Therefore, minimizing the imbalance helps to improve the efficiency and reliability of the power grid. Apart from earning profit for the energy load, QARG also reduces the grid imbalance. Note that the imbalances can be both positive and negative. Thus, we use mean squared imbalance as a metric in Fig. 6. In particular, we show the mean squared imbalance of the grid as a function of to different amounts of demand, i.e., E, when BI, QARG, Greedy, and HG schemes are applied, respectively. On average, the original mean squared imbalance without any regulation and the mean squared imbalance achieved by BI, QARG, the Greedy method, and HG are 161.2 MWh^2, 138.9 MWh^2,

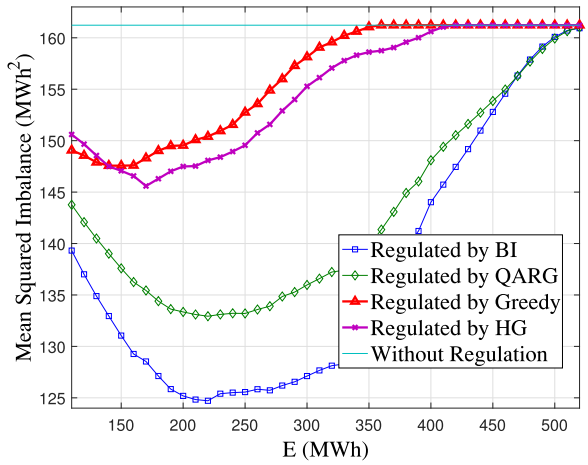


Fig. 6. Comparison of the achieved mean squared imbalance of regulated imbalance versus E.

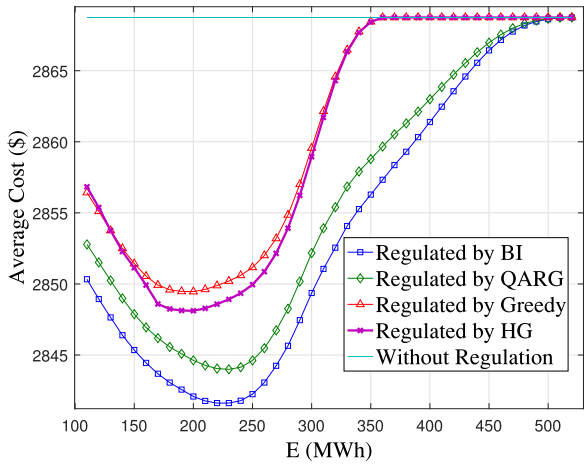


Fig. 7. Comparison of the bills of other utilities versus E.

144.0 MWh^2, 155.3 MWh^2, and 153.7 MWh^2 respectively. It can be seen that the mean squared imbalance of QARG and the three benchmark methods decrease with the demand when the demand is small and increase with the demand when the demand is large. It is because that when the demand is small, increased demand strengthens the ability to buy at off-load periods, and when the demand is large, this undermines the ability to sell at over-load periods. Meanwhile, the gaps between the mean squared imbalance obtained by QARG, the Greedy method, and HG increase with the demand. Recall that when the demand is large, without overall planning, the Greedy method and HG tend to buy energy at over-load period to fulfill the demand. In contrast, the QARG strategy reduces the grid imbalance with off-load buying and over-load selling.

Besides, QARG reduces not only the grid imbalance but also the the bills of the other utilities in the grid. In Fig. 7, we show the average bills of another utility in the same grid. For each demand, we evaluate the average bill over 90 days. In each hour, we assume there are 10 other utilities with zero-mean random imbalance in the grid. The realizations of the random imbalances are scaled such that the total imbalance is equal to the real-data.

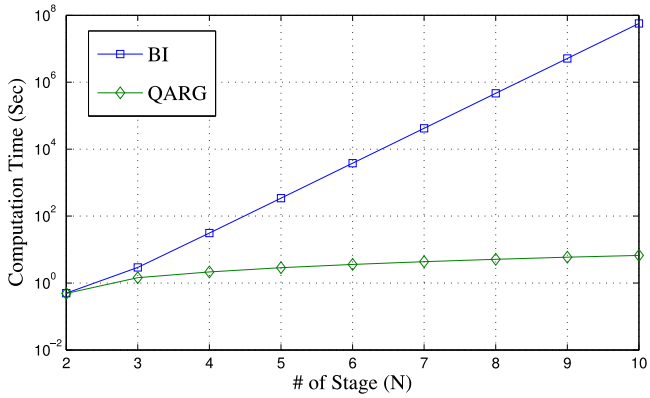


Fig. 8. Comparison of computational time versus E.

The bill is averaged over 100 imbalance realizations of particular utility. On average, the reduced bill achieved BI, QARG, the Greedy method, and HG are 677.9 \$, 592.3 \$, 375.8 \$, and 390.9 \$, respectively. The figure, together with Figs. 5 and 6, shows that QARG not only reduces the bill of the energy load itself but also the bills of other utilities and the imbalance of the local grid. Hence, it is a triple-win solution.

E. Computational Time Performance

In Fig. 8, we compare the computational time of QARG and the corresponding BI. For each number of stages, the computation is the average computational time of 90 days price realizations with 10 random demands for each day. In line with expectations, on one hand, the computational times of QARG is linearly increasing with the number of stages. On the other hand, the computational time of BI is exponentially increasing with the number of stages. This, together with Fig. 5, implies that QARG is a complexity-friendly and near-optimal strategy.

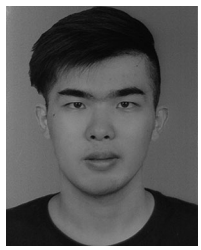
VI. Conclusion

In this paper, we proposed a profit-maximizing demand response scheme for an energy load in response to the real-time supply-demand imbalance in a real-time electricity market. We formulate the DR optimization problem into an MDP, and showed that the optimal solution possesses a threshold structure. Based on the threshold structure, we deduce a backward induction scheme to solve the MDP without the need of discretizing the state and action spaces. To further reduce the computational complexity, we proposed a QARG algorithm based on quadratic approximation and improved row generation strategies. Simulation results showed that QARG considerably outperforms the benchmark algorithms in terms of both the performance and complexity. Furthermore, it hits a triple win, in the sense that the real-time supply-demand imbalance is reduced, and the electricity bills of other market participants are reduced.

References

[1] A. L. Ott, "Experience with PJM market operation, system design, and implementation," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 528–534, May 2003.

[2] Y. J. Zhang, C. Zhao, W. Tang, and S. H. Low, "Profit maximizing planning and control of battery energy storage systems for primary frequency control," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 712–723, May 2018.
[3] Z. Liu, A. Wierman, Y. Chen, B. Razon, and N. Chen, "Data center demand response: Avoiding the coincident peak via workload shifting and local generation," *Performance Eval.*, vol. 70, no. 10, pp. 770–791, Oct. 2013.
[4] C. Eid, E. Koliou, M. Valles, J. Reneses, and R. Hakvoort, "Time-based pricing and electricity demand response: Existing barriers and next steps," *Utilities Policy*, vol. 40, pp. 15–25, Jun. 2016.
[5] L. P. Qian, Y. J. A. Zhang, J. Huang, and Y. Wu, "Demand response management via real-time electricity price control in smart grids," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1268–1280, Jul. 2013.
[6] R. Deng, Z. Yang, M. Y. Chow, and J. Chen, "A survey on demand response in smart grids: mathematical models and approaches," *IEEE Trans. Ind. Informat.*, vol. 11, no. 3, pp. 570–582, Jun. 2015.
[7] M. Pipattanasomporn, M. Kuzlu, S. Rahman, and Y. Teklu, "Load profiles of selected major household appliances and their demand response opportunities," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 742–750, Mar. 2014.
[8] T. H. Yu, S. Y. Kwon, K. M. Im, and J. H. Lim, "An RTP-based dimming control system for visual comfort enhancement and energy optimization," *Energy Buildings*, vol. 144, pp. 433–444, Jun. 2017.
[9] Q. Chen *et al.*, "Dynamic price vector formation model based automatic demand response strategy for PV-assisted EV charging station," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2903–2915, Apr. 2017.
[10] M. Parvania, M. Fotuhi-Firuzabad, and M. Shahidehpour, "ISO's optimal strategies for scheduling the hourly demand response in day-ahead markets," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2636–2645, Nov. 2014.
[11] M. Shafie-khah and J. P. S. Catalão, "A stochastic multi-Layer agent-based model to study electricity market participants behavior," *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 867–881, Mar. 2015.
[12] X. Fang, Q. Hu, F. Li, B. Wang, and Y. Li, "Coupon-based demand response considering wind power uncertainty: A strategic bidding model for load serving entities," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1025–1037, Mar. 2016.
[13] Z. Zhou, F. Zhao, and J. Wang, "Agent-based electricity market simulation with demand response from commercial buildings," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 580–588, Dec. 2011.
[14] K. S. Sedzro, A. J. Lamadrid, and M. C. Chuah, "Generalized minimax: A self-enforcing pricing scheme for load aggregators," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1953–1963, May 2018.
[15] H. Khani, M. R. Zadeh, and R. Seethapathy, "Optimal weekly usage of cryogenic energy storage in an open retail electricity market," in *Proc. IEEE Power Energy Soc. General Meet.*, Vancouver, BC, Canada, Jul. 2013, pp. 1–5.
[16] M. Shafie-khah, E. Heydarian-Forushani, M. E. H. Golshan, M. P. Moghaddam, M. K. Sheikh-El-Eslami, and J. P. S. Catalão, "Strategic offering for a price-maker wind power producer in oligopoly markets considering demand response exchange," *IEEE Trans. Ind. Informat.*, vol. 11, no. 6, pp. 1542–1553, Dec. 2015.
[17] M. Muratori and G. Rizzoni, "Residential demand response: Dynamic energy management and time-varying electricity pricing," in *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1108–1117, Mar. 2016.
[18] N. Vespermann, S. Delikaraoglou, and P. Pinson, "Offering strategy of a price-maker energy storage system in day-ahead and balancing markets," in *Proc. IEEE Manchester PowerTech*, Manchester, U.K., Jun. 2017, pp. 1–6.
[19] H. Mohsenian-Rad, "Coordinated price-maker operation of large energy storage units in nodal energy markets," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 786–797, Jan. 2016.
[20] M. Kohansal and H. Mohsenian-Rad, "Price-maker economic bidding in two-settlement pool-based markets: the case of time-shiftable loads," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 695–705, Jan. 2016.
[21] M. E. Lübbecke and J. Desrosiers, "Selected topics in column generation," *Oper. Res.*, vol. 53, no. 6, pp. 1007–1023, Dec. 2005.
[22] Y. Ye, D. Papadaskalopoulos, and G. Strbac, "Factoring flexible demand non-convexities in electricity markets," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2090–2099, Jul. 2015.
[23] W. Wei, J. Wang, and L. Wu, "Distribution optimal power flow with real-time price elasticity," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1097–1098, Jan. 2018.
[24] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY, USA: Wiley, 2014.
[25] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. II, 3rd ed. Belmont, MA, USA: Athena Scientific, 2007.
[26] M. Pishvaee, J. Razmi, and S. Torabi, "An accelerated benders decomposition algorithm for sustainable supply chain network design under uncertainty: A case study of medical needle and syringe supply chain," *Trans. Res. Part E: Logis. Transp. Rev.*, vol. 67, pp. 14–38, Jul. 2014.
[27] K. H. Rosen, *Handbook of Discrete and Combinatorial Mathematics*. Boca Raton, FL, USA: CRC Press, 1999.
[28] D. Li and X. Sun, *Nonlinear Integer Programming*, vol. 84. New York, NY, USA: Springer, 2006.
[29] "Historical market data," 2017. [Online]. Available: http://www.nordpool spot.com/historical-market-data/
[30] B. Neupane, T. B. Pedersen, and B. Thiesson, "Evaluating the value of flexibility in energy regulation markets," in *Proc. ACM e-Energy*, 2015, pp. 131–140.



Shuoyao Wang received the B.Eng. degree in information engineering from The Chinese University of Hong Kong, Hong Kong, in 2013. He is currently working toward the Ph.D. degree with The Chinese University of Hong Kong. His research interests include optimization theory, queueing analysis, dynamic programming, and reinforcement learning algorithm in EV charging scheduling and pricing, and electricity market operation.

Suzhi Bi (S'10–M'14) received the B.Eng. degree in communications engineering from Zhejiang University, Hangzhou, China, in 2009, and the Ph.D. degree in information engineering from The Chinese University of Hong Kong, Hong Kong, in 2013. From 2013 to 2015, he was a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. He was a Research Engineer Intern with the Institute for Infocomm Research, Singapore, in 2010, and a visiting student with the EDGE Laboratory, Princeton University, Princeton, NJ, USA, in 2012. Since 2015, he has been with the College of Information Engineering, Shenzhen University, Shenzhen, China, where he is currently an Associate Professor. His research interests include the optimizations in wireless information and power transfer, wireless medium access control, and smart power grid communications. He received the Shenzhen University Outstanding Young Faculty Award in 2015 and the IEEE SmartGridComm 2013 Best Paper Award. He is currently an Associate Editor of the IEEE ACCESS.

Ying-Jun Angela Zhang (S'00–M'05–SM'11) received the Ph.D. degree in electrical and electronic engineering from the Hong Kong University of Science and Technology, Hong Kong, in 2004. Since 2005, she has been with the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong, where she is currently an Associate Professor. Her research interests include mainly wireless communications systems and smart power systems, in particular optimization techniques for such systems. She is an Executive Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. She is also an Associate Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS. Previously, she served many years as an Associate Editor with the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, *Security and Communications Networks* (Wiley), and a Feature Topic in the IEEE COMMUNICATIONS MAGAZINE. She has served on the organizing committee of major IEEE conferences including ICC, GLOBECOM, SmartgridComm, VTC, CCNC, ICCC, MASS, etc. She is currently the Chair of IEEE ComSoc Emerging Technical Committee on Smart Grid. She was a Co-Chair of the IEEE ComSoc Multimedia Communications Technical Committee and the IEEE Communication Society GOLD Coordinator. She was the co-recipient of the 2014 IEEE ComSoc APB Outstanding Paper Award, the 2013 IEEE SmartgridComm Best Paper Award, and the 2011 IEEE Marconi Prize Paper Award on Wireless Communications. She was the recipient of the Young Researcher Award from The Chinese University of Hong Kong in 2011. As the only winner from engineering science, she has won the Hong Kong Young Scientist Award 2006, conferred by the Hong Kong Institution of Science. She is a Fellow of IET and a Distinguished Lecturer of IEEE ComSoc.