Modeling and Analysis of Baseline Manipulation in Demand Response Programs

Xiaochu Wang, *Student Member, IEEE* and Wenyuan Tang, *Member, IEEE*

***Abstract*—Baseline methods are used in demand response (DR) programs to estimate customers' intrinsic load so as to reward them properly. While the accuracy of baseline methods has drawn considerable attention, the strategic behavior regarding baseline manipulation has not been well explored in the literature. In this paper, we formulate the customer's payoff-maximizing problem as a Markov decision process (MDP). Several structural results have been established, including the characterization of underconsumption on event days and overconsumption on non-event days. We investigate the approximation of baseline methods to understand how the method parameters and the consumption statistics would affect the strategic behavior. Moreover, we develop a rollout algorithm, based on approximate dynamic programming, to solve the MDP efficiently. Finally, the proposed methodology is illustrated through case studies, which shed light on the analysis and design of baseline methods.**

***Index Terms*—Baseline manipulation, baseline method, demand response, dynamic programming, Markov decision process.**

I. Introduction

Demand response (DR) is considered as an effective and reliable solution to balance the power supply and demand when the power system is under stress [1]. In incentive-based DR programs, the program provider reaches an agreement with the customers, who are offered rebates for their active load adjustment according to the real-time instructions [2]. Baseline methods are critical in those DR programs, which estimate customers' intrinsic load so as to reward them properly. If the baseline is overestimated, the customers can be overpaid, which may lead to revenue deficits. On the other hand, underestimation of baselines may keep customers away from participation. Coming up with a perfect baseline method is demanding. First, the baseline is the counterfactual consumption that cannot be measured directly. In addition, the implementation of DR programs may motivate the customers to change their consumption patterns, which makes the baseline estimation more difficult [3].

Accuracy, integrity, and simplicity are three critical characteristics of practical baseline methods [4]. In particular, integrity means that a good baseline method is robust to customers' attempts to manipulate the system. A well-known class of baseline methods is the average-based method, which has been widely adopted by system operators and DR program providers such as PJM, NYISO, CAISO, and EnerNOC [5]. The baseline accuracy analysis has drawn considerable attention in the literature. The performance of several popular baseline methods is analyzed in [6] based on the bias metric. An elaborated error analysis of the existing baseline methods, including the average-based methods, exponential moving average, and regression methods, is provided in [7]. An economic analysis of a peak-time rebate program is also carried out. Their case study reveals that the utility pays at least half of its revenue as rebates solely due to the inaccuracy of baseline methods.

Although the inaccurate estimation of customer baseline is claimed as the main reason for the program inefficiency, baseline manipulation is also a key factor for this issue [8]. Customers are incentivized to inflate their baselines to gain higher rebates if the baseline methods were chosen improperly [9]. Such manipulation in a day-ahead load response program is illustrated in [10]. Consider a particular average-based method, High$X$of$Y$, which picks the highest X consumption measurements from the last Y days preceding the DR event, and calculates the average of the X measurements as the customer's baseline. In that case, customers may intentionally inflate their baselines by overconsuming energy on non-event days, in exchange for a possibly high rebate on a future event day. For example, a stadium turn on the lights during the daytime to increase the probability of getting a large amount of rebates [11].

Baseline manipulation has not been investigated in the literature until recently. The customer's manipulation is confirmed as a rational behavior based on a two-stage stochastic programming formulation [12]. In reality, the customer's current consumption may affect the rebates several days later, not just tomorrow. To investigate the customer's manipulation behavior, an infinite-horizon model is developed in [13], but only applicable to High$Y$of$Y$, a special case of High$X$of$Y$ when $X = Y$. A multi-stage dynamic programming model is also developed under the High$Y$of$Y$ method [14], in which DR events are modeled as independent Bernoulli random variables, without the consideration of their correlations. We believe that the state transition satisfies the Markov property and thus formulate a Markov decision process (MDP) to solve the customer's optimization problem. Apart from the manipulation analysis of the existing baseline methods, incentive-compatible mechanisms have been developed to eliminate the strategic behavior [15], [16], possibly at the cost of higher rebates and complexity. A novel contract is proposed in [17] to address the manipulation through self-report baselines, but that procedure may hinder customers' participation and is not compatible with the current framework of DR programs.

Manuscript received November 26, 2020; revised May 11, 2021; accepted June 21, 2021. *(Corresponding author: Wenyuan Tang.)*

The authors are with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695 USA (e-mail: xwang78@ncsu.edu; wtang8@ncsu.edu).

Our major contributions of this paper are summarized in the following:

1) A Markov decision process is formulated to model the customer's payoff-maximizing problem, which takes into account the correlations between DR events.
2) Several structural results are established for the general class of average-based methods, which provide insights into the strategic behavior of baseline manipulation.
3) The approximation of baseline methods is investigated to understand how the method parameters and the consumption statistics would affect the strategic behavior.
4) A rollout algorithm, based on approximate dynamic programming, is developed to solve the MDP efficiently, and the case studies shed light on the analysis and design of baseline methods.

The rest of this paper is organized as follows. In Section II, we introduce the MDP formulation of the customer's problem. Several structural results are established in Section III. We then investigate the approximation of baseline methods in Section IV. To solve the MDP efficiently, we develop a rollout algorithm in Section V, followed by the case study in Section VI. Section VII concludes the paper.

II. Problem Formulation

The MDP is a mathematical framework for solving sequential decision-making problems under uncertainties. Rational customers who aim to maximize the expected payoffs continuously adapt their actions to the underlying stochastic process of DR events, which can be modeled by the MDP formulation.

A. DR Event

A DR event occurs when the system reliability is under stress, typically associated with excess demand and high wholesale electricity prices. A day is called an event day if there is a DR event, and a non-event day otherwise. Let $\boldsymbol{y}_t$ summarize the exogenous circumstances such as grid and weather conditions on day t, and $\mathcal{E}$ be the set of conditions that trigger a DR event. Then $\mathbf{1}_{\mathcal{E}}(\boldsymbol{y}_t)$ indicates whether day t is an event day:

$$\mathbf{1}_{\mathcal{E}}(\boldsymbol{y}_t) = \begin{cases} 1, & \boldsymbol{y}_t \in \mathcal{E}, \\ 0, & \boldsymbol{y}_t \notin \mathcal{E}. \end{cases} \tag{1}$$

The characterization of $\mathcal{E}$ will be detailed in the case studies.

B. Baseline Method

For simplicity, we consider a particular hour (typically a peak hour when a DR event is likely to occur) for all days. For day t, let $\boldsymbol{x}_t = (x_{t,1}, \dots, x_{t,Y})$ be the consumption on the most recent Y non-event days (with $x_{t,1}$ the most recent). A baseline method is a function of the relevant recent consumption, denoted by $B : \mathbb{R}^Y \to \mathbb{R}$, with the interpretation that the value of the function is an estimate of the intrinsic load. The well-known class of average-based baseline methods include High$X$of$Y$, Low$X$of$Y$, and Mid$X$of$Y$. The High$X$of$Y$ baseline of day t is defined as the average of the X highest measurements among $x_{t,1}, \dots, x_{t,Y}$:

$$H_Y^X(\boldsymbol{x}_t) = \left(\frac{1}{X}\right) \max_{1 \le i_1 < \dots < i_X \le Y} (x_{t,i_1} + \dots + x_{t,i_X}). \tag{2}$$

As a special case (where $X = Y$), the High$Y$of$Y$ baseline is simply the average of $x_{t,1}, \dots, x_{t,Y}$. Similarly, the Low$X$of$Y$ baseline of day t is defined as the average of the X lowest measurements among $x_{t,1}, \dots, x_{t,Y}$:

$$L_Y^X(\boldsymbol{x}_t) = \left(\frac{1}{X}\right) \min_{1 \le i_1 < \dots < i_X \le Y} (x_{t,i_1} + \dots + x_{t,i_X}). \tag{3}$$

The Mid$X$of$Y$ (where X and Y have the same parity) baseline of day t is defined as the average of the X middle measurements among $x_{t,1}, \dots, x_{t,Y}$:

$$M_Y^X(\boldsymbol{x}_t) = \left(\frac{1}{X}\right) \left(x_{t,(\frac{Y-X}{2}+1)} + \dots + x_{t,(\frac{Y+X}{2})} \right). \tag{4}$$

where $x_{t,(1)}, \dots, x_{t,(Y)}$ is a permutation of $x_{t,1}, \dots, x_{t,Y}$ such that $x_{t,(1)} \le \dots \le x_{t,(Y)}$.

Among those average-based baseline methods, High$X$of$Y$ has been widely adopted in practice:

- PJM: High4of5 for weekdays; High2of3 for weekends.
- NYISO: High5of10 for weekdays; High2of3 for weekends.
- CAISO: High10of10 for weekdays; High4of4 for weekends.

For that reason, we will focus on the High$X$of$Y$ method in this paper. The same methodology can be applied to Low$X$of$Y$ and Mid$X$of$Y$ as well.

C. Customer's Utility Function

A customer obtains benefits from electricity consumption, captured by the utility function $u_t(a_t; z_t)$, where a_t is the customer's electricity consumption and z_t is an external parameter indicating the benefit variation on different days. The utility function is typically non-decreasing and concave in consumption [18]. There are several forms of utility functions in the literature, e.g., the quadratic and exponential utility functions. Our model is valid under any non-decreasing and concave utility functions.

Let $c_t(a_t)$ be the cost of purchasing electricity from load serving entities. Specifically, we consider the flat retail price ω so that $c_t(a_t) = \omega a_t$. The net utility function is given by

$$\pi_t(a_t; z_t) = u_t(a_t; z_t) - c_t(a_t). \tag{5}$$

In the absence of DR programs, a rational customer would choose the intrinsic baseline a_t^{B} as the optimal consumption to maximize the net utility:

$$a_t^{\mathrm{B}} \in \arg\max_{0 \le a_t \le \hat{a}} \pi_t(a_t; z_t), \tag{6}$$

where $\hat{a}$ is a constant consumption limit.

D. Markov Decision Process

Consider a finite time horizon of T days, indexed by $t = 0, 1, \ldots, T-1$, with a dummy terminal day $t = T$. The state at day t is given by $\boldsymbol{s}_t = (\boldsymbol{x}_t, \boldsymbol{y}_t)$. The decision variable at day t is the consumption a_t, subject to a physical constraint $0 \leq a_t \leq \hat{a}$. The transition from $\boldsymbol{x}_t$ to $\boldsymbol{x}_{t+1} = (x_{t+1,1}, \ldots, x_{t+1,Y})$ is deterministic, given by

$$\boldsymbol{x}_{t+1} = \begin{cases} \boldsymbol{x}_t = (x_{t,1}, \ldots, x_{t,Y}), & \boldsymbol{y}_t \in \mathcal{E}, \\ (a_t, x_{t,1}, \ldots, x_{t,Y-1}), & \boldsymbol{y}_t \notin \mathcal{E}, \end{cases} \tag{7}$$

whereas $\boldsymbol{y}_{t+1}$ is independent of $(\boldsymbol{x}_t, a_t)$ conditioned on y_t, with the transition probability density function denoted by $p(\boldsymbol{y}_{t+1}|\boldsymbol{y}_t)$. The immediate payoff function is given by

$$g_t(\boldsymbol{s}_t, a_t) = \pi_t(a_t; z_t) + \mathbf{1}_{\mathcal{E}}(\boldsymbol{y}_t) r (B(\boldsymbol{x}_t) - a_t)_+ \tag{8}$$

for all $t < T$, with $g_T(\boldsymbol{x}_T, \boldsymbol{y}_T) = 0$. Here r is the rebate price per unit of load reduction $(B(\boldsymbol{x}_t) - a_t)_+$, where $(x)_+ = \max\{x, 0\}$.

The customer's problem is to find an optimal policy $\{a_t^*(\boldsymbol{s}_t), \forall t\}$ to maximize the expected payoff

$$\mathrm{E}_{\{\boldsymbol{y}_t, \forall t\}} \left[\sum_{t=0}^{T} g_t(\boldsymbol{s}_t, a_t) \right]. \tag{9}$$

Following a dynamic programming approach, we obtain the Bellman equation:

$$V_t(\boldsymbol{s}_t) = \max_{0 \leq a_t \leq \hat{a}} \{ g_t(\boldsymbol{s}_t, a_t) + \mathrm{E}_{\boldsymbol{y}_{t+1}|\boldsymbol{y}_t}[V_{t+1}(\boldsymbol{s}_{t+1})] \} \tag{10}$$

for all $t < T$, with $V_T(\boldsymbol{s}_T) = 0$. The optimal policy is given by

$$a_t^*(\boldsymbol{s}_t) \in \arg\max_{0 \leq a_t \leq \hat{a}} \{ g_t(\boldsymbol{s}_t, a_t) + \mathrm{E}_{\boldsymbol{y}_{t+1}|\boldsymbol{y}_t}[V_{t+1}(\boldsymbol{s}_{t+1})] \} \tag{11}$$

for all $t < T$.

III. Properties of the Optimal Policy

We apply dynamic programming to investigate the properties of the optimal policy, which provide insights into the strategic behavior of baseline manipulation.

A. Properties of Baseline Methods

The monotonicity of the average-based methods can be easily seen.

Proposition 1. *For fixed X and Y, $H_Y^X(\boldsymbol{x}_t)$, $L_Y^X(\boldsymbol{x}_t)$, and $M_Y^X(\boldsymbol{x}_t)$ are increasing in $\boldsymbol{x}_t$. For fixed Y and $\boldsymbol{x}_t$, $H_Y^X(\boldsymbol{x}_t)$ is decreasing in X, and $L_Y^X(\boldsymbol{x}_t)$ is increasing in X.*

The convexity of the High$X$of$Y$ and Low$X$of$Y$ methods can be shown from the fact that pointwise maximum preserves convexity.

Proposition 2. *For fixed X and Y, $H_Y^X(\boldsymbol{x}_t)$ is convex in $\boldsymbol{x}_t$, and $L_Y^X(\boldsymbol{x}_t)$ is concave in $\boldsymbol{x}_t$.*

The concept of supermodularity will also be useful for deriving the structural results.

Definition 1 (Supermodularity)**.** *For any $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{x}' = (x_1', \ldots, x_n')$ in $\mathbb{R}^n$, define their* join *as $\boldsymbol{x} \vee \boldsymbol{x}' = (\max\{x_1, x_1'\}, \ldots, \max\{x_n, x_n'\})$ and their* meet *as $\boldsymbol{x} \wedge \boldsymbol{x}' = (\min\{x_1, x_1'\}, \ldots, \min\{x_n, x_n'\})$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is* supermodular *if for any $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^n$,*

$$f(\boldsymbol{x}) + f(\boldsymbol{x}') \leq f(\boldsymbol{x} \vee \boldsymbol{x}') + f(\boldsymbol{x} \wedge \boldsymbol{x}'). \tag{12}$$

f is submodular *if $(-f)$ is supermodular.*

The supermodularity of the High$X$of$Y$ and Low$X$of$Y$ methods can be derived from the definition.

Proposition 3. *For fixed X and Y, $H_Y^X(\boldsymbol{x}_t)$ is submodular in $\boldsymbol{x}_t$, and $L_Y^X(\boldsymbol{x}_t)$ is supermodular in $\boldsymbol{x}_t$.*

B. Structural Results of the Optimal Policy

The following lemma is intuitive, which states that the higher the recent consumption (and hence the higher baseline), the higher the optimal value function.

Lemma 1. *For any given $H_Y^X(\boldsymbol{x}_t)$, $L_Y^X(\boldsymbol{x}_t)$, or $M_Y^X(\boldsymbol{x}_t)$, $V_t(\boldsymbol{s}_t) = V_t(\boldsymbol{x}_t, \boldsymbol{y}_t)$ is increasing in $\boldsymbol{x}_t$.*

Proof. Since $B(\boldsymbol{x}_t)$ is increasing in $\boldsymbol{x}_t$ by Proposition 1, it can be easily seen that the rebate received on day t,

$$\mathbf{1}_{\mathcal{E}}(\boldsymbol{y}_t) r (B(\boldsymbol{x}_t) - a_t)_+, \tag{13}$$

is also increasing in $\boldsymbol{x}_t$. Then in light of the Bellman equation (10), it can be proved by backward induction that $V_t(\boldsymbol{s}_t) = V_t(\boldsymbol{x}_t, \boldsymbol{y}_t)$ is increasing in $\boldsymbol{x}_t$. □

The following result states that compared to the intrinsic baseline without strategizing, the optimal decision is to underconsume on event days, and to overconsume on non-event days.

Theorem 1. *The intrinsic baseline a_t^B as defined in (6) and the optimal policy $a_t^*(\boldsymbol{s}_t)$ as defined in (11) have the following relationship:*

$$\begin{cases} a_t^*(\boldsymbol{x}_t, \boldsymbol{y}_t) \leq a_t^B, & \boldsymbol{y}_t \in \mathcal{E}, \\ a_t^*(\boldsymbol{x}_t, \boldsymbol{y}_t) \geq a_t^B, & \boldsymbol{y}_t \notin \mathcal{E}. \end{cases} \tag{14}$$

The proof of this theorem, as well as the proofs for theorems 2, 3, 4, can be found in [19]. While the DR program helps alleviate the supply shortage on event days, it also encourages excess consumption on non-event days, resulting in an overall loss of social welfare. This fundamental result provides the rationale behind baseline manipulation.

Next, we characterize the optimal policy on event days.

Theorem 2. *Let a_t^B be the intrinsic baseline as defined in (6). Define*

$$a_t^U \in \arg\max_{0 \leq a_t \leq \hat{a}} \{ \pi_t(a_t; z_t) + r(B(\boldsymbol{x}_t) - a_t) \}, \tag{15}$$

which satisfies $a_t^U \leq a_t^B$. Define

$$B_t^{TH} = \frac{\pi_t(a_t^B; z_t) - \pi_t(a_t^U; z_t)}{r} + a_t^U. \tag{16}$$

The optimal policy on event days is a threshold policy with respect to the current baseline $B(\boldsymbol{x}_t)$:

$$a_t^*(\boldsymbol{x}_t, \boldsymbol{y}_t) = \begin{cases} a_t^B, & B(\boldsymbol{x}_t) \le B_t^{TH}, \\ a_t^U, & B(\boldsymbol{x}_t) > B_t^{TH}, \end{cases} \tag{17}$$

for all $\boldsymbol{y}_t \in \mathcal{E}$.

The above theorem shows that a rational customer will reduce their consumption only if the current baseline exceeds some threshold. Otherwise, the optimal decision remains the intrinsic consumption.

The following lemmas give support to the derivation of structural results for the optimal policy on non-event days.

Lemma 2 (Supermodularity composition rule)**.** *If a function $f : \mathbb{R} \to \mathbb{R}$ is convex and increasing (decreasing) and another function $g : \mathbb{R}^n \to \mathbb{R}$ is monotonic and supermodular (submodular), then $f(g(\boldsymbol{x}))$ in supermodular in $\boldsymbol{x} \in \mathbb{R}^n$.*

The proof of this lemma can be found in [20]. The following lemma is helpful in establishing comparative static results when the objective function is not differentiable [21].

Lemma 3 (Topkis's theorem)**.** *If f is supermodular (submodular) in $(\boldsymbol{x}, \vartheta)$ and $\mathcal{G}$ is a lattice, then $\boldsymbol{x}^*(\vartheta) \in \arg\max_{\boldsymbol{x}\in\mathcal{G}} f(\boldsymbol{x}, \vartheta)$ is increasing (decreasing) in ϑ.*

Finally, we characterize the optimal policy on non-event days.

Theorem 3. *If the baseline method $B(\boldsymbol{x}_t)$ is supermodular in $\boldsymbol{x}_t$, then the optimal policy on non-event days, $a_t^*(\boldsymbol{x}_t, \boldsymbol{y}_t)$ for $\boldsymbol{y}_t \notin \mathcal{E}$, is increasing in $\boldsymbol{x}_t$.*

According to Proposition 3, the above theorem applies for the Low$X$of$Y$ baseline (including High$Y$of$Y$ as a special case), in which the customer would consume more on non-event days when the recent consumption is higher. For the High$X$of$Y$ method, such structural results are not readily available, which motivates the approximation of baseline methods to enable useful characterization of the optimal policy.

IV. Approximation of Baseline Methods

The impact from program parameters, i.e., X and Y, on the optimal policies is little-known under the average-based methods. In addition, how the statistical distribution of the customer's consumption affects the calculated baseline is also unclear. To address these issues, we propose a methodology for approximating the baseline methods, based on which more structural results will be established.

A. Methodology

The baseline approximation consists of two steps: (i) a linearization in terms of the sample mean and extreme order values; (ii) an estimation of the extreme order values. The main idea of the linearization is to use the weighted sample mean and extreme order values to approximate the average-based methods, where the program parameters, i.e., X and Y, are used to construct the corresponding weights.

Given the recent consumption $\boldsymbol{x}_t = (x_{t,1}, \dots, x_{t,Y})$, we have the sample mean $\bar{x}_t = (1/Y)(x_{t,1} + \cdots + x_{t,Y})$ and the sample maximum $x_{t,\max} = \max\{x_{t,1}, \dots, x_{t,Y}\}$. The approximated High$XofY$ is defined as the convex combination of the sample mean and the sample maximum:

$$\hat{H}_Y^X(\boldsymbol{x}_t) = \lambda \bar{x}_t + (1-\lambda) x_{t,\max}, \tag{18}$$

where $\lambda = \frac{X-1}{Y-1} \in [0,1]$ is the weight of the sample mean. Considering that $\hat{H}_Y^X(\boldsymbol{x}_t)$ also includes X and Y, we do not show them inside the parentheses to simplify the expression. It can be seen that when $X = 1$ or $X = Y$, we have $\hat{H}_Y^X(\boldsymbol{x}_t) = H_Y^X(\boldsymbol{x}_t)$. Similarly, after introducing the sample minimum $x_{t,\min} = \min\{x_{t,1}, \dots, x_{t,Y}\}$ and the sample median $x_{t,\text{med}}$, one can define the approximated Low$X$of$Y$ and Mid$X$of$Y$ as $\hat{L}_Y^X(\boldsymbol{x}_t) = \lambda \bar{x}_t + (1-\lambda) x_{t,\min}$ and $\hat{M}_Y^X(\boldsymbol{x}_t) = \lambda \bar{x} + (1-\lambda) x_{t,\text{med}}$, respectively.

The above linearization approximates the average-based methods using only the sample mean and extreme order values. However, the sample maximum $x_{t,\max}$ in (18) still prevents us from deriving structural results since $x_{t,\max}$ is essentially equivalent to High1of$Y$. Therefore, we need to come up with an approximation of the sample maximum $x_{t,\max}$. We adopt the approach in [22], which is used to estimate the sample maximum under normal distributions. Suppose the standard normal distribution $\mathcal{N}(0,1)$ is truncated by the interval $[\alpha, \beta]$, denoted by $\mathcal{N}(0,1,\alpha,\beta)$. Then the sample maximum can be estimated as

$$f(Y) = \Phi_t^{-1}[(0.5 + \zeta_t)^{1/Y}], \tag{19}$$

where Y is the sample size, $\Phi_t(\cdot)$ is the cumulative distribution function of the truncated normal distribution, and ζ_t is a calculated parameter. Suppose that the sample is drawn from a truncated normal distribution with a mean μ and a standard deviation σ. Then the sample maximum can be estimated in the form of $\mu + f(Y)\sigma$ [22]. However, the customer's consumption, especially under baseline manipulation, does not necessarily follow a (truncated) normal distribution. To address this issue, we utilize the form $\mu + f(Y)\sigma$ and construct the sample maximum estimation as follows:

$$\hat{x}_{t,\max} = \bar{x}_t + f(Y) s_t, \tag{20}$$

where $\bar{x}_t$ is the sample mean, s_t is the sample standard deviation with $s_t^2 = \frac{1}{Y-1}\sum_{i=1}^{Y}(x_{t,i} - \bar{x}_t)^2$, and $f(Y)$ is an increasing function of the sample size Y (the larger the sample size, the larger the sample maximum). Without the assumption of the customer consumption distribution, we do not have an explicit expression of $f(Y)$ like (19). Instead, $f(Y)$ can be fitted based on the actual consumption data. Then, based on (18), one can obtain the approximated High$X$of$Y$ method:

$$\hat{H}_Y^X(\boldsymbol{x}_t) = \bar{x}_t + \frac{Y-X}{Y-1} f(Y) s_t, \ \forall Y \ge 2. \tag{21}$$

Note that the fitted $f(Y)$ is not required to derive the following structural results. The theoretical analysis of the validity of this approximation will be shown later.

B. Structural Results under the Approximation

1) The monotonicity of the $\hat{H}_Y^X(\boldsymbol{x}_t)$: We show the monotonicity of the approximated baseline in its parameters in the following proposition.

Proposition 4. *The $\hat{H}_Y^X(\boldsymbol{x}_t)$ is increasing in $\boldsymbol{x}_t$, but decreasing in X.*

The monotonicity of $\hat{H}_Y^X(\boldsymbol{x}_t)$ in Y is underdetermined but relies on the selection of X. We give two special selections of X. (i) $X = 1 : \hat{H}_Y^X = \bar{x}_t + f(Y)s_t$ is increasing in Y. (ii) $X = Y : \hat{H}_Y^X = \bar{x}_t$ is independent of Y. Besides, the monotonicity of $\hat{H}_Y^X(\boldsymbol{x}_t)$ in s_t shows that the more scattered the customer consumption, the higher the approximated baseline.

2) The customer's manipulation: We first show the submodularity of $\hat{H}_Y^X(\boldsymbol{x}_t)$ and then present the structural result of the optimal policy.

Proposition 5. *The $\hat{H}_Y^X(\boldsymbol{x}_t)$ is submodular in $\boldsymbol{x}_t$.*

Proof. We have the $\hat{H}_Y^X(\boldsymbol{x}_t) = \bar{x}_t + \frac{Y-X}{Y-1} f(Y)s_t$, where the standard deviation can be written as

$$s_t(\boldsymbol{x}_t) = \sqrt{\frac{1}{Y-1}\left(\sum_{i=1}^{Y} x_{t,i}^2 - Y\bar{x}_t^2\right)}, \tag{22}$$

with $\bar{x}_t = \frac{1}{Y}(x_{t,1} + \cdots + x_{t,Y})$. The submodularity of $s_t(\boldsymbol{x}_t)$ can be verified via the composition rule from Lemma 2. □

The above proposition shows that the $\hat{H}_Y^X(\boldsymbol{x}_t)$ has the same submodularity as $H_Y^X(\boldsymbol{x}_t)$ in $\boldsymbol{x}_t$. Unfortunately, no structural results can be derived directly based on the submodularity. However, we can apply the same strategy used in Theorem 3 to establish the monotonicity of $a_t^*(\boldsymbol{x}_t, \boldsymbol{y}_t)$ for $\boldsymbol{y}_t \notin \mathcal{E}$ in the program parameter X. On a non-event day, the optimal policy is obtained by

$$a_t^*(\boldsymbol{x}_t, \boldsymbol{y}_t) \in \underset{0 \le a_t \le \hat{a}}{\arg\max}\{\pi_t(a_t) + \mathrm{E}_{\boldsymbol{y}_{t+1}|\boldsymbol{y}_t}[V_{t+1}(\boldsymbol{s}_{t+1})]\}, \tag{23}$$

for $\boldsymbol{y}_t \notin \mathcal{E}$. To show the monotonicity of $a_t^*(\boldsymbol{x}_t, \boldsymbol{y}_t)$ for $\boldsymbol{y}_t \notin \mathcal{E}$ in X, one only needs to show that the expectation term, or simply $\hat{H}_Y^X(\boldsymbol{x}_t)$, is supermodular in (a_t, X). We characterize the conditions that lead to the monotonicity in the following.

Theorem 4. *Given a fixed value of Y, the optimal policy on a non-event day, $a_t^*(\boldsymbol{x}_t, \boldsymbol{y}_t)$ for $\boldsymbol{y}_t \notin \mathcal{E}$, is increasing in X when the intrinsic consumption on that day is lower than the population average, and decreasing in X otherwise.*

The above theorem is illustrated in the following examples. On a low-consumption day, a rational customer inflates more under High10of10 compared to that under High1of10. This is reasonable because everyday matters under High10of10 but only one highest day counts for High1of10. On the other hand, a small X incentivizes the customer to inflate more on the high-consumption days. For example, under High1of10, a customer increases the load to secure a higher baseline for the future.

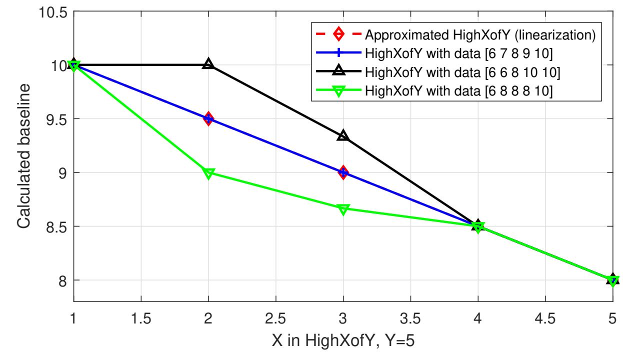


Figure 1. An illustration of the linearization error with three different data sets: {6, 7, 8, 9, 10}, {6, 6, 8, 10, 10}, and {6, 8, 8, 8, 10}. These three data sets share the same means and maximums, respectively.

C. Approximation Validity Analysis

Following the above structural results, a question then arises: does the approximation error affect the validity of the conclusions, i.e., Proposition 4 and Theorem 4? For Proposition 4, we observe that the monotonicity of High$X$of$Y$ in X is preserved in the final approximated High$X$of$Y$ (21). Thus, the monotonicity in Proposition 4 only depends on the approximation structure and is valid regardless of the approximation accuracy. For Theorem 4, it is derived based on the condition that $\hat{H}_Y^X(\boldsymbol{x}_t)$ is supermodular in (a_t, X). Considering that the sample maximum estimation (20) does not change the modularity, whether $\hat{H}_Y^X(\boldsymbol{x}_t)$ is supermodular in (a_t, X) depends on the accuracy of linearization. Here, we illustrate the linearization accuracy (error) in Fig. 1 based on three data sets. To show only the linearization error, these three data sets are selected so that their means and maximums are equal, respectively. One of them is uniformly distributed {6, 7, 8, 9, 10}, which leads to zero linearization error. For the other two cases, there are some approximation errors, especially when X is around $Y/2$. While it is intractable to quantify the impact of linearization accuracy on the supermodularity of $H_Y^X(\boldsymbol{x}_t)$ in (a_t, X), it is more likely that Theorem 4 applies to $H_Y^X(\boldsymbol{x}_t)$ as well when the approximation error is small enough.

V. Approximate Dynamic Programming via Rollout Algorithm

Solving the MDP problem using dynamic programming is computationally challenging due to the curse of dimensionality [23]. Therefore, approximate methods are required to efficiently solve the MDP problems with a large state space [24]. In this section, we apply the rollout algorithm to obtain a simulation-based forward dynamic programming solution.

1) The rollout algorithm: The key idea of the rollout algorithm is starting from some given heuristic and constructing another policy with better performance [25]. Essentially, the rollout algorithm is an online forward dynamic programming procedure that selects actions to obtain the maximum expected payoff calculated based on the given heuristic policy. The improved performance is guaranteed by the performance improvement properties [26], which we will show later. Rollout algorithms step forward in time and thus require estimating the payoff-to-go when evaluating decision rules. Specifically, given a heuristic policy, we calculate the expected payoff-to-go under the heuristic based on a Monte Carol simulation [27]. The formula in finding the optimal actions under the rollout algorithm is given as

$$a_t^{\mathrm{R}} \in \underset{0\le a_t\le \hat{a}}{\arg\max} \left\{ g_t(s_t,a_t) + \mathrm{E}\left[\sum_{i=t+1}^{T} g_i(s_i,a_i^{\mathrm{L}}(s_i))|s_t\right]\right\}, \tag{24}$$

where $a_t^{\mathrm{L}}(s_t)$ is the given heuristic policy. The sample paths in calculating the expected payoff-to-go can be generated by applying additive white Gaussian noise to customer historical consumption. The details of generating sample paths are given in the case study.

2) The heuristic policy: A heuristic policy is any method to select decision rules within the decision rule space. In this paper, a linear policy is selected as the heuristic for the rollout algorithm. Linear policies with k features can be expressed as $a(s;\boldsymbol{\theta}) = \theta_1\cdot\phi_1(s)+\cdots+\theta_k\cdot\phi_k(s)$, where $\phi_1(s),\ldots,\phi_k(s)$ are the extracted features and $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_k)$ is the corresponding weight vector. Given the optimal policy on event days in (17), we approximate the optimal policy on non-event days based on one feature, given by

$$a_t^{\mathrm{L}}(\boldsymbol{x}_t,\boldsymbol{y}_t) = a_t^{\mathrm{B}} + \theta_t\phi(\boldsymbol{s}_t), \tag{25}$$

for $\boldsymbol{y}_t \notin \mathcal{E}$, where the feature $\phi(\boldsymbol{s}_t)$ can be the maximum of the recent consumption. The idea behind this linear policy is to provide customers an easy way to inflate their baseline based on the recent consumption. The best parameter θ_t^* can be obtained by maximizing the expected payoff from the current stage to the end stage, i.e.,

$$\theta_t^* \in \underset{0\le\theta_t\le1}{\arg\max}\, \mathrm{E}\left[\sum_{i=t}^{T} g_i(s_i,a_i^{\mathrm{L}}(s_i))|s_t\right], \tag{26}$$

where the linear policy $a_t^{\mathrm{L}}(s_t)$ on non-event days and event days are given in (25) and (17). For simplicity, a stationary linear policy $\theta^* = \theta_{t=1}^*$ is applied for simulation.

3) Performance improvement property [26]: The performance improvement property of the rollout algorithm is illustrated in the following definition.

Definition 2 (Rollout Improvement Property)**.** *For a heuristic a_t^L and a rollout policy a_t^R, we say that a_t^R is rollout improving if for $t = 1,2,\ldots,T$,*

$$\mathrm{E}\left[\sum_{i=t}^{T} g_i(s_i,a_i^L(s_i))|s_t\right] \le \mathrm{E}\left[\sum_{i=t}^{T} g_i(s_i,a_i^R(s_i))|s_t\right]. \tag{27}$$

This rollout improvement property is assured based on the construction of a_t^{R} in (24) and a_t^{L} in (26).

VI. Case Study

In this section, we conduct a case study to show the customer's manipulation levels under the High$X$of$Y$ methods. Due to the curse of dimensionality, we only give the results using dynamic programming with $Y \le 7$. Cases with larger Ys are conducted via the rollout algorithm.

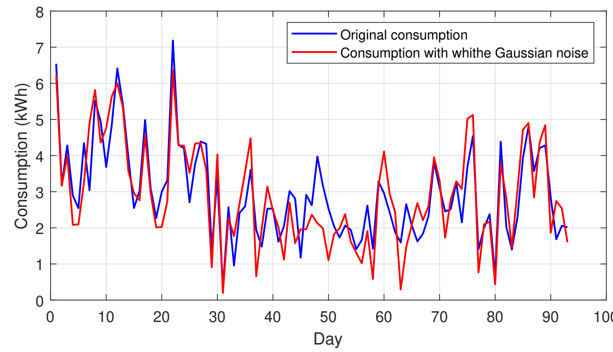


Figure 2. An illustration of the initial consumption data and the data with additive white Gaussian noise. Days are selected from high-consumption seasons, mostly in spring and winter.

A. Bias and Data Set

We introduce the metric of bias to evaluate the performance of the baseline method. Let $\{a_t^{\mathrm{B}}\}_{t\in\mathbb{H}}$ be the customer's historical consumption at the DR-event hour and $\mathcal{D}$ be the set of the event days. Then we can define the bias as

$$\frac{\sum_{t\in\mathcal{D}}\left[B(\boldsymbol{x}_t) - a_t^{\mathrm{B}}\right]}{\sum_{t\in\mathcal{D}} a_t^{\mathrm{B}}} \times 100\%, \tag{28}$$

where a_t^{B} is the intrinsic consumption and $B(\boldsymbol{x}_t)$ is the calculated baseline. By comparing the biases in the base case (without DR programs) with the manipulation case (with DR programs), one can identify the customer baseline manipulation level.

To properly define the referential bias, we first introduce the customer's historical consumption $\{a_t^{\mathrm{B}}\}_{t\in\mathbb{H}}$ that comes from the *UMass* smart home data set [28]. This data set consists of load and weather data for 114 households in a 15-minute interval for two years and three months. Here, one data set from a typical customer[1] (resident #21) in 2015 is used for illustration. We use the hourly consumption on the weekdays and remove the data on holidays. We set 9:00 am as the DR hour (with the highest average consumption). Furthermore, a total of 93 days from peak consumption seasons is considered for evaluation. However, the limited data set cannot serve as a stable reference in evaluating baseline manipulation. To address this issue, we apply additive white Gaussian noise to the initial data set and create large data samples. As a result, the time-series information of the initial data set is preserved and the generated large data set provides a stable reference. The initial data and the data with additive white Gaussian noise are illustrated in Fig. 2. The signal-to-noise ratio is set as 3dB to balance the time-series information and the consumption randomness. Note that the lower bound is set to zero to avoid negative consumption.

In addition, we generate independent DR event sequences to increase the randomness of the sample path, which is a $T = 93$ days consumption and DR signals. A total of $N = 100$ paths are generated for simulation where the DR events are activated according to exogenous circumstances.

[1]We choose the customers based on two criteria: (i) without bad data and (ii) showing a morning peak.

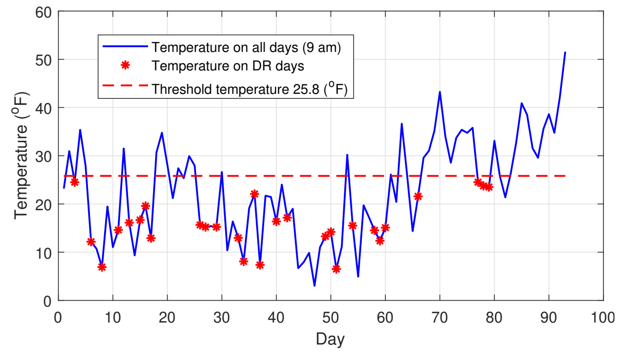


Figure 3. The DR event generation based on a threshold temperature.

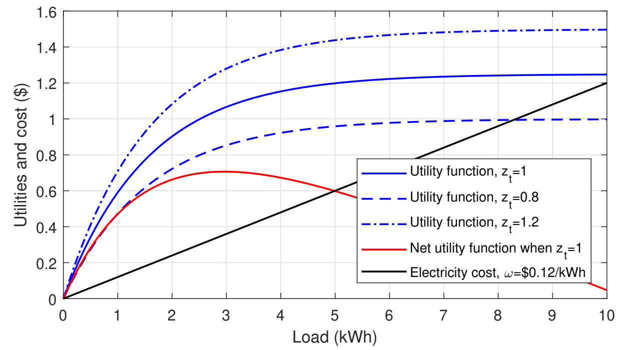


Figure 4. The customer's utility functions and net utility function.

The temperature is generally considered as the main factor among those circumstances. For instance, a logistic regression model is formulated to learn the DR event probability versus daily peak temperature [14]. Inspired by this, we adopt an empirical approach to generate the DR events: a fixed DR event probability is assigned for days once the temperature drops to a certain threshold temperature (for winter) and zero probability otherwise. An illustration of this strategy is given in Fig. 3, where the temperature data is from the 93 days at 9:00 am. We set the threshold temperature as 25.8 ^oF and the fixed probability as 50%, resulting in an overall DR rate around 30%. More DR event sequences are generated by adding white Gaussian noise to the temperature data. Note that the above approach is considered as the actual DR generation mechanism in this paper. Actually, the customers do not have the perfect information about DR activities and need to estimate this uncertainty. In our model formulation, we use $p(\boldsymbol{y}_{t+1}|\boldsymbol{y}_t)$ to capture the transition of the exogenous circumstances. However, this setting complicates the customer's payoff-maximization problem because she needs to estimate both the transition probability and the DR generation mechanism. A practical approach is to directly estimate the transition probabilities $p_0 := p(\boldsymbol{y}_{t+1} \in \mathcal{E}|\boldsymbol{y}_t \notin \mathcal{E})$ and $p_1 := p(\boldsymbol{y}_{t+1} \in \mathcal{E}|\boldsymbol{y}_t \in \mathcal{E})$ from historical DR event data. This is reasonable since a time series analysis of the temperature data shows a high correlation of temperatures in two neighboring days. For the historical DR event data, we consider the result from Fig. 3. As a result, the transition probability is estimated as: $p_0 = 0.25$ and $p_1 = 0.38$.

B. Estimation of Customer's Utility Function

Given a customer's historical consumption $\{a_t^{\mathrm{B}}\}_{t\in\mathbb{H}}$, we propose an innovative way to estimate the customer utility function. The consumption mean $\bar{a}$ and upper bound $\check{a}$ of the historical consumption are used for this estimation. We define the utility function as $u_t(a_t; z_t) = z_t \cdot u(a_t)$, where $u(a_t)$ is the referential utility under which the consumption mean $\bar{a}$ is the customer's optimal decision. Besides, we use the exponential form of the utility function as $u(a_t) = \gamma(1 - e^{-a_t/\rho})$, where γ and ρ are two parameters called utility ratio and utility shape, respectively. The determination of these two parameters is given as follows.

The utility shape ρ: we introduce the maximum relative utility as $\check{u} := \frac{u_t(\check{a},z_t)}{\lim_{a_t\to\infty} u_t(a_t;z_t)} = 1 - e^{-\check{a}/\rho}$. In practice, we choose $\check{u} = 0.99$[2]. Therefore, the utility shape parameter ρ can be determined as

$$\rho = -\frac{\check{a}}{\ln(1-\check{u})}. \tag{29}$$

The utility ratio γ: γ is defined as the ratio of one-unit utility to one monetary value ($). According to the definition of $u(a_t)$, i.e., under which the average consumption $\bar{a}$ is optimal, we have $\bar{a} \in \arg\max_{0\le a_t \le \hat{a}} [u(a_t) - c_t(a_t)]$. Based on the first-order optimality condition, the derivative is zero at $\bar{a}$, i.e., $u'(\bar{a}) := (\gamma/\rho)e^{-\bar{a}/\rho} = c_t'(\bar{a}) = \omega$. Therefore, we have the utility ratio as

$$\gamma = \omega\rho e^{\frac{\bar{a}}{\rho}}. \tag{30}$$

Based on the historical consumption data from the customer #21, we have the utility parameters as $\rho = 1.56$ and $\gamma = 1.25$. Fig. 4 illustrates the customer utilities under $z_t = [0.8, 1, 1.2]$, shown as the blue curves. The net utility function, shown as the red curve, is concave in consumption.

C. Manipulation Results via Dynamic Programming

We solve the MDP problem by following the dynamic programming approach shown in (10). To find the optimal policies, we discretize the action into 10 values. Due to the curse of dimensionality, we only consider cases with $Y \le 7$. Note that the following simulation results are based on a single typical customer (resident #21) with 100 simulated paths described in Section VI.A.

1) Bias and manipulation in terms of X in High$X$of$Y$: Fig. 5 shows the results for High$X$of$Y$ when $Y = 7$. The bias is used to evaluate the performance of baseline methods, which is defined in (28). The manipulation level is the difference between the biases with and without demand response. We first observe that the bias is decreasing in X, which matches the monotonicity of High$X$of$Y$ in X. The bias is around zero when $X = 7$ because the DR events are independently generated of customer consumption. The bias with DR manipulation is much higher than the counterpart without DR. The difference between the bias with and without DR is captured by the manipulation curve. We also notice that the manipulation is a single-peaked function in X, where $X = 6$ and 7 leads to the lowest manipulation level.

[2] This is a subjective choice and any value that is close to 1 is reasonable. Furthermore, the results are not very sensitive to the selection of $\check{u}$.

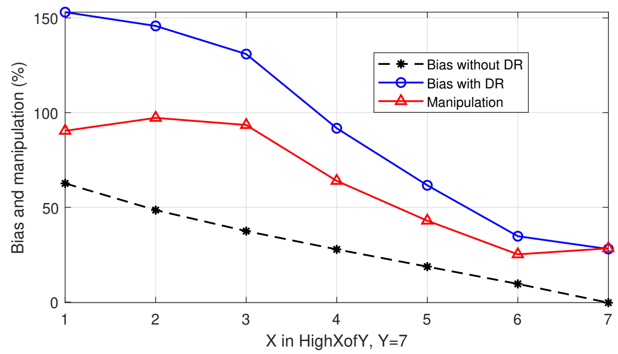


Figure 5. The bias and manipulation in terms of X when $Y = 7$ and $r = \$0.12$/kWh.

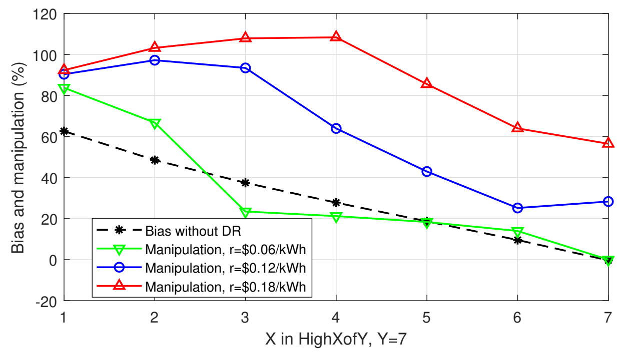


Figure 6. The bias and manipulation in terms of r and X when $Y = 7$.

2) Results in terms of rebate price r: We compare the results under three rebate prices, i.e., \$0.06/kWh, \$0.12/kWh, and \$0.18/kWh, with the retail price as \$0.12/kWh. The curves in Fig. 6 show that the higher the rebate price, the more the customer manipulation. The three curves also preserve the single-peaked property. When the rebate price is \$0.06/kWh, the manipulation is relatively low when $X \geq 3$ and almost zero when $X = 7$. This shows that a low rebate price and a large X can effectively alleviate the customer's manipulation.

3) Results under High$X$of$Y$ with different values of Y: Fig. 7 shows the manipulation under the High$X$of$Y$ method when $Y = 3, \ldots, 7$. All the manipulation curves preserve the single-peaked property and the shapes look like a tilted and stretched letter **Z** when $Y \geq 5$. The manipulation level almost remains the same under the High$Y$of$Y$ method when $Y = 3, \ldots, 7$.

D. Evaluation of Approximation Error

To better show the validity of our baseline approximation, we evaluate the approximation (linearization) error here. The data used for evaluation is the consumption with and without manipulation. We set $Y = 7$ for evaluation and choose the manipulated consumption data under $X = 2$, which shows the most manipulation according to Fig. 7. The box plot of the approximation errors is given in Fig. 8 where more detailed information is specified in the caption. The average absolute errors are small in the without manipulation case, with the maximum as 5.9%. In the manipulation case, the average absolute errors have a maximum value of 15.7%. However,

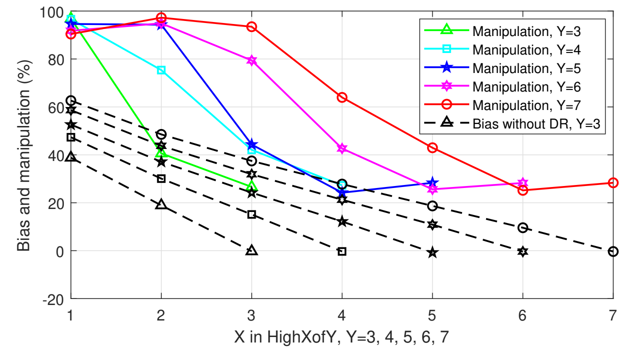


Figure 7. The bias and manipulation for High$X$of$Y$, $Y = 3, \ldots, 7$, $r = \$0.12$/kWh.

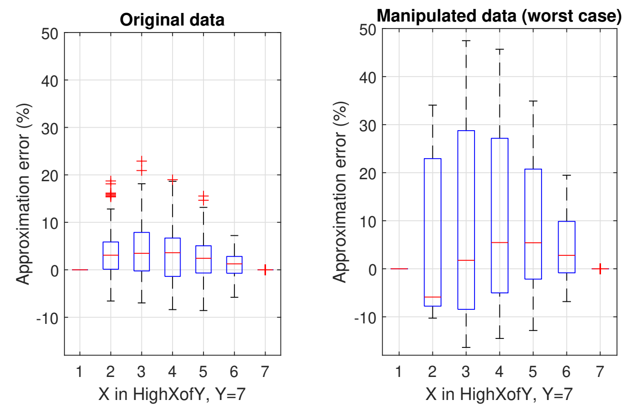


Figure 8. Evaluation of approximation errors based on data with and without manipulation. The instance with most manipulation ($X = 2$) is selected as the manipulated case. The average absolute errors for the above two cases are [0, 5%, 5.9%, 5.7%, 4.4%, 2.5%, 0] and [0, 12.9%, 15.7%, 15.3%, 12.0%, 6.7%, 0] for the without and with manipulation cases, respectively.

the median and the mean (calculated as 9.9%) are below 10% even for the worst scenario. Together with Fig. 1, we believe this amount of approximation error will not change the supermodularity of H_Y^X in $(\boldsymbol{x}_t, X)$ and thus the conclusion is valid.

E. Manipulation Results via Rollout Algorithm

In this subsection, we show the results obtained via the linear policy-based rollout algorithm and compare them with the exact solutions. The simulation is based on the single typical customer (resident #21) and the generated simulated paths. We generate 10 paths to calculate the expected payoff-to-go for the linear policy, where the best parameter is

$$\theta^*(X) = \begin{cases} 0.148, & \text{when } X = 1, \ldots, 7, \\ 0.125, & \text{when } X = 8, 9, 10. \end{cases}$$

When $Y = 7$, we have the same threshold parameters, with $\theta^*(X) = 0.148$ for $X = 1, \ldots, 5$, and $\theta^*(X) = 0.125$ for $X = 6, 7$. Note that the rollout algorithm is a simulation-based method. A total of 1000 paths are generated to obtain the rollout policy (24). Fig. 9 shows the simulation results for High$X$of$Y$ when $Y = 7$ and 10. The two blue curves are the manipulation levels obtained via the rollout algorithm

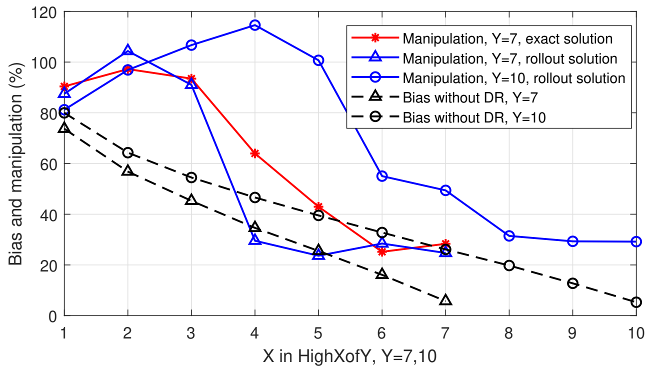


Figure 9. Comparison between the exact solution and the rollout solution for High$X$of$Y$, $Y = 7$ and 10, $r =$ \$0.12/kWh.

while the red curve shows the exact result for $Y = 7$. The two approximated curves preserve the single-peaked property. The manipulation levels drop quickly around $X = Y/2$ and remain almost the same when X is close to Y. We also compare the results from dynamic programming and from the rollout algorithm when $Y = 7$: although there are some differences in manipulation when $X \leq 5$, these two curves are similar in shape and almost match when $X = 6$ and 7. Considering that we mainly use the rollout algorithm to obtain a fast approximated solution to the manipulation trend analysis (qualitative analysis), this rollout algorithm is considered effective.

VII. Conclusion

In this paper, we give the baseline manipulation analysis of the used accepted average baseline methods. We establish a Markov decision process to formulate the customer's profit-maximization problem. By comparing the optimal policies to the intrinsic baseline, we verify a rational customer's behavior: underconsumption on event days and overconsumption on non-event days. An approximated High$X$of$Y$ method is proposed to understand how the distribution of customer consumption and the selection of program parameters affect the calculated baseline. The case study shows that the customer baseline manipulation is a single-peaked function in X and remains almost the same under different High$Y$of$Y$ methods. These results give practical information for DR program providers to design accurate and robust baseline methods.

References

[1] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A survey on demand response programs in smart grids: pricing methods and optimization algorithms," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 152–178, 2015.

[2] P. Palensky and D. Dietrich, "Demand side management: demand response, intelligent energy systems, and smart loads," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 3, pp. 381–388, 2011.

[3] H.-p. Chao, "Demand response in wholesale electricity markets: the choice of customer baseline," *Journal of Regulatory Economics*, vol. 39, no. 1, pp. 68–88, 2011.

[4] N. Rossetto, "Measuring the intangible: an overview of the methodologies for calculating customer baseline load in PJM," 2018.

[5] EnerNOC, "The demand response baseline," Tech. Rep., 2009.

[6] T. K. Wijaya, M. Vasirani, and K. Aberer, "When bias matters: An economic assessment of demand response baselines for residential customers," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1755–1763, 2014.

[7] S. Mohajeryami, M. Doostan, A. Asadinejad, and P. Schwarz, "Error analysis of customer baseline load (CBL) calculation methods for residential customers," *IEEE Transactions on Industry Applications*, vol. 53, no. 1, pp. 5–14, 2017.

[8] F. A. Wolak, "Residential customer response to real-time pricing: The anaheim critical peak pricing experiment," 2007.

[9] X. Wang and W. Tang, "To overconsume or underconsume: Baseline manipulation in demand response programs," in *2018 North American Power Symposium (NAPS)*. IEEE, 2018, pp. 1–6.

[10] "Filing of changes to day-ahead load response program," ISO New England Docket No. ER08-538-000, Tech. Rep., 2008.

[11] S. Borenstein, "Peak-time rebates: money for nothing?" https://www.greentechmedia.com/articles/read/peak-time-rebates-money-for-nothing, 2014, [Online; accessed 15-March-2018].

[12] J. Vuelvas and F. Ruiz, "Demand response: Understanding the rational behavior of consumers in a peak time rebate program," in *2015 IEEE 2nd Colombian Conference on Automatic Control (CCAC)*. IEEE, 2015, pp. 1–6.

[13] H.-p. Chao and M. DePillis, "Incentive effects of paying demand response in wholesale electricity markets," *Journal of Regulatory Economics*, vol. 43, no. 3, pp. 265–283, 2013.

[14] D. Ellman and Y. Xiao, "Incentives to manipulate demand response baselines with uncertain event schedules," *IEEE Transactions on Smart Grid*, 2020.

[15] D. P. Zhou, M. Balandat, M. A. Dahleh, and C. J. Tomlin, "Eliciting private user information for residential demand response," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 189–195.

[16] Y. Li and N. Li, "Mechanism design for reliability in demand response with uncertainty," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 3400–3405.

[17] J. Vuelvas, F. Ruiz, and G. Gruosso, "Limiting gaming opportunities on incentive-based demand response programs," *Applied Energy*, vol. 225, pp. 668–681, 2018.

[18] P. Samadi, A.-H. Mohsenian-Rad, R. Schober, V. W. Wong, and J. Jatskevich, "Optimal real-time pricing algorithm based on utility maximization for smart grid," in *2010 First IEEE International Conference on Smart Grid Communications*. IEEE, 2010, pp. 415–420.

[19] X. Wang and W. Tang, "Analysis and evaluation of baseline manipulation in demand response programs," *arXiv preprint arXiv:2011.10681*, 2020.

[20] D. Simchi-Levi, X. Chen, J. Bramel *et al.*, "The logic of logistics," *Algorithms, and Applications for Logistics and Supply Chain Management, Theory*, 2005.

[21] D. M. Topkis, "Minimizing a submodular function on a lattice," *Operations research*, vol. 26, no. 2, pp. 305–321, 1978.

[22] C.-C. Chen and C. W. Tyler, "Accurate approximation to the extreme order statistics of gaussian samples," *Communications in Statistics-Simulation and Computation*, vol. 28, no. 1, pp. 177–188, 1999.

[23] W. B. Powell, *Approximate Dynamic Programming: Solving the curses of dimensionality*. John Wiley & Sons, 2007, vol. 703.

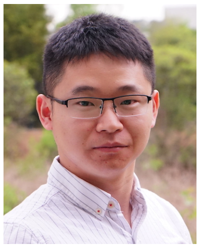
[24] D. P. Bertsekas and D. A. Castanon, "Rollout algorithms for stochastic scheduling problems," *Journal of Heuristics*, vol. 5, no. 1, pp. 89–108, 1999.

[25] D. P. Bertsekas, "Rollout algorithms for discrete optimization: A survey," *Handbook of combinatorial optimization*, pp. 2989–3013, 2013.

[26] J. C. Goodson, B. W. Thomas, and J. W. Ohlmann, "A rollout algorithm framework for heuristic solutions to finite-horizon stochastic dynamic programs," *European Journal of Operational Research*, vol. 258, no. 1, pp. 216–229, 2017.

[27] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-dynamic programming*. Athena Scientific Belmont, MA, 1996, vol. 5.

[28] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, J. Albrecht *et al.*, "Smart*: An open data set and tools for enabling research in sustainable homes," *SustKDD, August*, vol. 111, no. 112, p. 108, 2012.



Xiaochu Wang (S'21) received the B.Eng. degree in automation from Chongqing University, Chongqing, China, in 2013, the M.S. degree in control science and engineering from Shanghai Jiao Tong University, Shanghai, China, in 2016. He is currently pursuing the Ph.D. degree in electrical engineering at North Carolina State University, Raleigh, NC, USA.

His research interests include power system operation, optimization and economics, and electricity markets with a focus on demand response programs.

Wenyuan Tang (S'14–M'17) received the B.Eng. degree in electrical engineering from Tsinghua University, Beijing, China, in 2008, the M.S. degree in electrical engineering, the M.A. degree in applied mathematics, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 2010, 2014, and 2015, respectively.

He was a Postdoctoral Scholar with the University of California, Berkeley, and Stanford University. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, USA. His research interests include electricity markets, energy data analytics and machine learning, and control and optimization for power systems.