# Contract-Based Demand Response Mechanism for Commercial and Industrial Customers

Sajad Parvizi, Mina Montazeri, and Hamed Kebriaei, *Senior Member, IEEE*

***Abstract*— Designing an optimal Demand Response (DR) program necessitates having information on the utility of users. However, that is not easily accessible since the users may not be willing to report their personal information. In this paper, we propose an incentive-based DR mechanism for Commercial and Industrial (C&I) customers, leveraging contract theory to address the challenges of incomplete information and ensure the participation of customers in the DR program. Designing a DR mechanism includes specifying an incentive reward, penalty, and the corresponding demand reduction as functions of the customer's utility information. These functions are obtained through an optimization problem that maximizes the grid operator's profit, ensures truthful reporting of personal information and participation of the customers in the DR program. After formulating the optimal mechanism, the main technical challenge is reformulating the obtained non-convex and computationally inefficient optimization problem to a tractable and convex one. The extensive numerical evaluations demonstrate the effectiveness of our proposed framework in comparison to existing benchmark schemes.**

***Note to Practitioners*—Imbalances between supply and demand, particularly during the hot summer months, can result in issues such as power outages. This problem is detrimental for both grid operators and customers alike. DR programs have been implemented in numerous countries to address these challenges. C&I customers, due to the schedulable nature of their loads, make for ideal targets for these programs. However, the main obstacle to implementing such programs is customer participation. Incentive rewards offered in these programs for load reduction are fixed for all customers, discouraging those with low rewards from participating and imposing high costs on grid operators for those with high rewards. On the other side, designing unique rewards and penalties for each customer necessitates information about their preferences, which grid operators cannot access. In this paper, we propose an incentive mechanism based on the contract between the grid operator and C&I customers that ensures customers truthfully report their preferences while simultaneously maximizing grid operator profit and guaranteeing the participation of all customers. The proposed DR program can be implemented for large non-residential customers in order to reduce the peak load of the grid.**

***Index Terms*— Incentive-based demand response, contract theory, incomplete information, dual decomposition.**

Manuscript received 10 February 2024; accepted 5 April 2024. This article was recommended for publication by Associate Editor W. Chen and Editor X. Xie upon evaluation of the reviewers' comments. The work of Hamed Kebriaei was supported in part by the Institute for Research in Fundamental Sciences (IPM) under Grant CS 1402-4-208. *(Corresponding author: Hamed Kebriaei.)*

Sajad Parvizi is with the School of ECE, College of Engineering, University of Tehran, Tehran 1417935840, Iran (e-mail: sajadparvizi8305@gmail.com).

Mina Montazeri is with the School of ECE, College of Engineering, University of Tehran, Tehran 1417935840, Iran, and also with the Urban Energy Systems Laboratory, Swiss Federal Laboratories for Materials Science and Technology, 8600 Dübendorf, Switzerland (e-mail: mina.montazeri@empa.ch).

Hamed Kebriaei is with the School of ECE, College of Engineering, University of Tehran, Tehran 1417935840, Iran, and also with the School of Computer Science, Institute for Research in Fundamental Sciences (IPM), Tehran 1668836471, Iran (e-mail: kebriaei@ut.ac.ir).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TASE.2024.3389999.

Digital Object Identifier 10.1109/TASE.2024.3389999

## I. Introduction

In recent years, Demand Response (DR) has emerged as a pivotal strategy in the electricity industry, addressing the pressing challenge of balancing electricity supply and demand while simultaneously promoting energy efficiency and reducing emissions [1]. By incentivizing and empowering customers to modify their energy consumption patterns, DR programs have become indispensable in bridging the gap between supply and demand dynamics [2]. Among the diverse customer segments, Commercial and Industrial (C&I) customers, due to having schedulable loads, are often considered more suitable for DR programs. Despite their comparatively smaller numbers, these customers possess considerable energy consumption capacities and have the potential to deliver substantial responses, making them highly desirable candidates for effective DR implementation [3]. Thus, conducting a comprehensive examination of the behavior and characteristics of C&I customers becomes paramount in successfully implementing DR strategies and achieving their intended goals.

DR programs are divided into two broad categories: incentive-based DR and price-based DR [4]. Price-based DR relies on time-varying electricity prices to motivate customers to adjust their consumption patterns, but its limited dispatchability and potential financial risks have hindered its widespread adoption [5]. In response to these challenges, incentive-based DR has emerged as an alternative approach that offers greater flexibility and reduced financial risks for customers. Under incentive-based DR, customers are rewarded with incentives for reducing their electricity consumption.

Most of the current incentive-based DR structures enable C&I customers to participate in the day-ahead market through bilateral contracts. Traditionally, these incentives were proposed to be constant and solely based on customer participation in the program. For instance, the California Public Utilities Commission (CPUC) introduced an innovative program in 2021 called the Emergency Load Reduction Program (ELRP), which aims to prevent rotating outages during peak summer electricity usage periods. The ELRP utilizes an incentive-driven DR approach, whereby commercial customers temporarily curtail their electricity usage to mitigate strain on the power grid and receive a payment 2$ for each unit of demand reduction accomplished [6]. Similarly, the Center Point Energy Commercial Load Management Program in Texas provides financial incentives to large C&I customers who successfully reduce their energy consumption during high-demand periods. Participants receive financial incentives (up to 31.50 $ per kW) for their participation, which increase based on the amount of load reduction achieved [7].

Constant incentive schemes may not adequately meet the specific requirements of different industries, as the financial impact of curtailing or shifting energy consumption varies among various industries. Therefore, to compensate for the lack of financial losses, the DR incentives should be offered based on the characteristics of the target industries individually [8]. Designing personalized incentives is therefore crucial to encourage the participation of all customers. Another critical challenge in determining the optimal DR program pertains to the requirement of detailed information on the preferences of customers. While existing literature assumes that customers' utility functions are available to the grid operator or truthfully revealed by the customers themselves [9], this assumption may not hold in real-world scenarios [10]. Contract theory is a promising solution to address the aforementioned challenges. The basic idea of the contract theory is to offer an appropriate contract to the agents, such that all agents are encouraged to reveal their personal information truthfully, while the designer maximizes his profit [11].

This paper introduces an incentive-based DR program tailored for C&I customers, drawing upon contract theory. This mechanism guarantees participation and truthful behavior of C&I customers in the DR program under incomplete information (i.e., the grid operator does not have access to the actual personal information of a particular customer and only knows the probability distribution over all customers' types). The grid operator designs contracts that incorporate incentive reward, penalty, and demand reduction based on the reported personal information provided by the customer in the DR event. Additionally, it is crucial for the grid operator to account for load constraints imposed by the customers themselves, as well as their own operational requirements, when determining the optimal levels of demand reduction, incentive rewards, and penalties. The penalties are designed so that customers do not deviate from the contracted items. This consideration encompasses the diverse range of C&I customer loads, including shiftable, critical, and curtailable loads. The optimization problem associated with designing the contract with a multi-load scenario can be technically intricate, particularly when addressing shiftable loads, which introduces a coupled optimization challenge. To overcome this complexity, we employ the dual decomposition method, as outlined in [12]. To demonstrate the effectiveness of our proposed framework, we compare our proposed scheme's performance with the existing benchmark scheme through extensive numerical evaluation. The main contributions of the paper are summarized as follows:

- We use contract theory to propose an incentive mechanism for DR programs for C&I customers under incomplete information from the utility function of the customers.
- Given that the optimization problem arising from contract theory is inherently non-convex, we employ a mathematical reformulation technique to transform it into a convex optimization problem with a reduced set of constraints.
- We establish the necessary and sufficient conditions for the feasibility and optimality of the proposed mechanism, ensuring that it simultaneously maximizes the profit of the grid operator, and guarantees individual rationality, and incentive compatibility.
- We consider all kinds of C&I customer loads, including curtailable, critic, and shiftable loads, which impose coupling constraints to the optimization problem. To overcome this challenge, the dual decomposition method has been incorporated.

## II. Literature Review

In terms of incomplete information, works [13], [14] used an "auction mechanism" to deal with users' unknown information in resource allocation problems, which is different from our task allocation problem. To address the challenges posed by incomplete information in task allocation, a significant body of research has adopted a "mechanism design" approach to incentivize customers to provide truthful information to aggregators. For instance, in [15], based on the baseline information and marginal utility reported by customers, they present a mechanism that the probability of choosing customers and their reward for reducing their consumption depends on the information reported by other customers. They will be penalized if they are not selected and do not consume according to the reported baseline. The authors in [16] propose a mechanism based on a self-reported baseline, randomized selection of consumers, and penalty for uninstructed deviations to ensure that the baseline inflation is controlled. Similarly, in [17] DR contract between a user and an aggregator is developed through the probability of call. In these works, the penalty function in combination with randomized calling allows the designer to control the truthful behavior of customers. However, due to randomized calling, it is possible that customers will not participate in the event and will not receive a profit despite the truthful report of their private information.

The authors in [18] proposed a Vickrey-Clarke Groves (VCG) mechanism for demand-side management applications to encourage users to report their demand information truthfully. In [19], an incentive compatible mechanism based on VCG auction is proposed for energy and reserve market clearing in multi-area power systems. Authors in [20], opt for a VCG-like approach to achieve social welfare maximization, but they omit the direct-revelation approach of the typical VCG mechanism. This mechanism guarantees truthful user participation and preserves users' privacy. It is important to note that although these studies focus on designing truthful mechanisms for social welfare maximization, our study differs in that our goal is the profit maximization of the mechanism designer, such as the grid operator. As VCG mechanisms are not effective for profit maximization [11], alternative approaches are necessary.

In this context, researchers have turned to the contract theory approach that deals with the problem of profit maximization in the presence of incomplete information [21], [22]. The contract-based demand response management problem is formulated in [21] as a maximization problem of the electricity market's utility, while incentivizing the prosumers to purchase an optimal personalized amount of electricity at the announced price. Similarly, in [22] formulates and solves contract-based optimization problems between the microgrid operator and the sellers or buyers to determine the rewards and amount of sold/purchased energy. References [21] and [22] assume a linear form for the incentive function based on the allocation value. This limitation hampers the ability to capture non-linear relationships or more complex incentive structures within the allocation value. Reference [23] proposed a data-driven contract approach to model energy trading; however, the possibility of non-fulfillment of the contracted quantities by the customers was not considered. References [24] and [25] consider this issue in the model of the energy trading problem. However, the penalty rate for all types of customers was considered a fixed value.

Different from the existing works, in this paper, we propose a contract mechanism that simultaneously includes allocation, reward, and penalty with the purpose of 1-profit maximization of the designer 2-participation of all C&I customers (individual rationality constraints) 3-truthful behavior of C&I customers in the DR program (incentive compatibility constraints) 4-compliance with the necessary constraints of grid operator and C&I customers. Penalty is also a part of contracts to avoid deviation of customers after the contract and designed according to their type. In addition, due to the consideration of different kinds of customers' loads, there is a coupling between the allocations of the contracts, which makes solving the optimization problem different from existing works and to tackle this we use the dual decomposition method.

## III. System Model and Problem Formulation

Our proposed incentive-based demand response scheme's system model is depicted in Figure 1 and involves a Grid Operator (GO) and multiple C&I customers (CUs). The grid operator draws up a contract offering demand reduction alongside corresponding incentive rewards and penalties. This contract's design is particularly crucial when there is a prediction of a significant increase in electricity demand or a shortage of supply, which motivates triggering a DR event to maintain grid stability and prevent blackouts. C&I customers have the choice to participate in this contract, aiming to maximize their incentive reward income while considering the cost of dissatisfaction due to demand reduction and penalties for contract deviation. On event days identified by the grid operator using available information, a forward contract is designed with C&I customers to minimize overall electricity consumption. This is a strategic measure to ensure a balance between supply and demand. The optimal outcome of the DR program necessitates that the grid operator has complete information regarding the customers' dissatisfaction functions. Nonetheless, acquiring such information is a challenge, as customers might untruthfully report their personal information to gain more profit, acting as rational and selfish players. To mitigate this issue and enable the design of a more effective and efficient

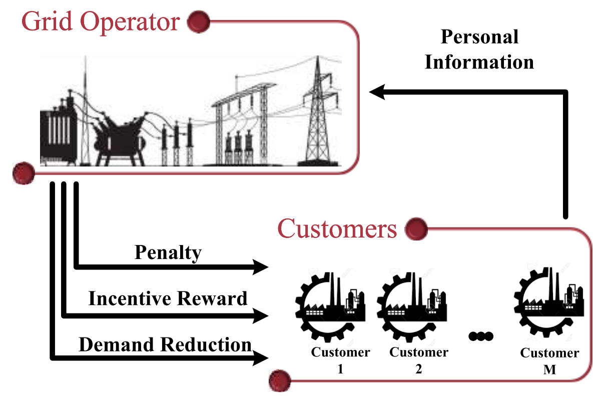



Fig. 1. An interactive framework for proposed incentive-based DR scheme in smart grid.

DR program, customers are encouraged to voluntarily report their personal information, or 'type'. To address the potential misinformation and encourage truthful reporting of personal information, the grid operator employs contract theory. This theory is used to establish a contract that delineates the optimal incentive rewards, demand reduction targets, and penalties, ensuring that customers not only participate in the DR program but also report their personal information truthfully.

### A. Customer Model

We consider a set of C&I customers that interact with the grid operator. To characterize the heterogeneity of customers, we categorize them into $N$ classes based on their personal information, referring to them as type-$i$ customers where $i$ belongs to the set $i \in \mathcal{N} = \{1, 2, \ldots, N\}$. A typical customer is equipped with an energy management system and a range of equipment that can be categorized into three load categories [26] including:

*Critic loads ($d_{cr}$):* This kind of load must be immediately met, as they operate continuously and cannot be rescheduled. These loads require a stable and reliable power supply to ensure uninterrupted operation, as any disruption or outage can lead to significant financial losses (e.g., data center). *Shiftable loads ($d_{sh}$):* These loads are programmable and can be moved to other times. However, it needs to be met at the end of the day (e.g., maintenance activities). *Curtailable loads ($d_{cu}$):* This kind of load can be potentially reduced by sacrificing a certain level of satisfaction (e.g., compressed air systems, lightning).

A customer with type-$i$ has a hourly predicted load $\hat{d}^t_{n,i} = \hat{d}^t_{cu,i} + \hat{d}^t_{sh,i} + \hat{d}^t_{cr,i}, t \in \mathcal{T} = \{1, 2, \ldots, 24\}$ that includes all these three types of loads. To ensure a minimum level of customer satisfaction, the following constraints are imposed on the level of demand reduction [27]:

$$\begin{cases} 0 < D^t_i \leq K^t_r \quad \forall t \in \mathcal{T}, \ \forall i \in \mathcal{N}, & (1) \\ \sum_{t \in \mathcal{T}} D^t_i \leq K_s \quad \forall i \in \mathcal{N}, & (2) \end{cases}$$

where $D^t_i$ is the amount of demand reduction offered to type-$i$ customer in time slot $t$, $K^t_r = \min_m\{\hat{d}^t_{n,m} - \hat{d}^t_{cr,m}\}$ and $K_s = \min_m\{\sum_{t \in \mathcal{T}} \hat{d}^t_{cu,m}\} \forall t \in \mathcal{T}$. Note that constraint (1) ensures that the critic load is satisfied, which means to guarantee that each customer has at least as much power as its critical load, the demand reduction should fall within the range of 0 and $\hat{d}_n^t - \hat{d}_{cr}^t$ [27]. Also, since the critical load must be satisfied and cannot be reduced, ensuring the satisfaction of the shiftable load during the day in constraint (2) necessitates that the total demand reduction is, at most, the size of the other available load, i.e., the sum of curtailable loads ($\sum_{t\in\mathcal{T}} D_i^t \leq \sum_{t\in\mathcal{T}} \hat{d}_{cu,i}^t$) [27]. Due to the grid operator's lack of knowledge regarding the actual type of each customer, the worst-case scenario is considered for both constraints.

*1) Utility Function:* The grid operator provides customers with an incentive reward in response to their demand reduction efforts. However, the demand reduction can cause discomfort for the customers, and such discomfort is commonly modeled as a dissatisfaction cost. Furthermore, due to unforeseen circumstances that may occur, potential CUs may not be able to provide the contracted demand reduction during the DR event. For this reason, the grid operator sets a penalty for CUs, proportional to the deviation from their contracted demand reduction. Consequently, the type-$i$ customer's utility consists of three components: reward term, dissatisfaction term, and penalty term, which can be expressed as follows [15] and [28]:

$$U_{CU,i}^t(D_i^t, \lambda_i^t, P_i^t) = \tilde{D}_i^t \lambda_i^t - \alpha_i (\tilde{D}_i^t)^2 - (D_i^t - \tilde{D}_i^t) P_i^t, \quad (3)$$

where $\lambda_i^t$ is the incentive reward that offered to type-$i$ customer and $P_i^t$ is the penalty rate that type-i customer incurs to deviate from the demand reduction contracted. During the DR event, each customer fulfills a percentage of the demand reduction according to their circumstances and performs their preferred demand reduction $\tilde{D}_i^t$. Thus, CUs are rewarded for demand reduction performed and penalized for demand reduction shortfalls below their contracted demand reduction. The term $\alpha_i(\tilde{D}_i^t)^2$ is the dissatisfaction cost that type-$i$ customer incurs to reduce their demand. We define the demand reduction valuation coefficient $\alpha_i > 0$ as the type of customer that represents a particular valuation of the demand reduction. A higher valuation coefficient indicates that customers are less incentivized to reduce their demand. Moreover, the type of customer is considered as its personal information (i.e., only known to the customer itself, and neither the grid operator nor other customers know the type). Here, we consider that the customers' types belong to a discrete and finite space [24], [29] and without loss of generality, it is assumed that $0 < \alpha_1 < \alpha_2 < \ldots < \alpha_N$. The following Theorem suggests a penalty function in which the optimal demand reduction for a consumer becomes equal to the contracted demand reduction.

*Theorem 1: The penalty function $P_i^t = 2\alpha_i D_i^t - \lambda_i^t$ avoids of type-i customer to deviate from the contracted demand reduction, i.e., $\tilde{D}_i^t * = D_i^t$.*

*Proof:* The preferred demand reduction $\tilde{D}_i^t$ which maximizes the utility of the customer is

$$\tilde{D}_i^{t\star} = \arg\max_{\tilde{D}_i^t} U_{CU,i}(D_i^t, \lambda_i^t, P_i^t). \quad (4)$$

From first order optimality condition we get,

$$\lambda_i^t - 2\alpha \tilde{D}_i^t + P_i^t = 0. \quad (5)$$

Clearly the second-order condition is also met and thus:

$$\tilde{D}_i^{t\star} = \frac{\lambda^t + P^t}{2\alpha_i}. \quad (6)$$

Hence by choosing $P_i^t = 2\alpha_i D_i^t - \lambda_i^t$ we have $\tilde{D}_i^{t\star} = D_i^t$, and the proof is completed. □

*Remark 1: Although the penalty function depends on the personal information of type-i customer, nevertheless, we show that the contract is designed in such a way that the customer has no incentive to misreport its type to the GO.*

### B. Grid Operator Model

The grid operator tries to reduce the cost of capacity procurement. At each time slot, the grid operator faces an expected demand deficit denoted as $D_{req}^t$. To address this deficit, the grid operator purchases demand reductions from customers. This can be expressed as follows [20] and [28]:

$$\sum_{i=1}^{N} M f_i D_i^t = D_{req}^t \quad \forall t \in \mathcal{T}, \quad (7)$$

where $f_i$ is The probability that a customer belongs to type-$i$, and $M$ is total number of customers. The variable $Mf_i$ represents the count of customers with type-$i$. As mentioned before, the type of each customer is part of its personal information, and grid operator has no information about it. However, we assume that grid operator knows a common probability distribution over all customers' types based on available statistical information or online learning [24]. The utility of the grid operator is evaluated based on its net profit, which is derived from subtracting the payments made to customers as part of the contract from the gross profit generated by demand reduction. For simplicity, we assume that the gross profit made due to the reduction of demand $D_i^t$ is directly proportional to a factor of $D_i^t$. Consequently, the utility of the grid operator is defined as follows:

$$U_{GO}(\mathbf{D}, \boldsymbol{\lambda}) = \sum_{i=1}^{N}\sum_{t=1}^{T} M f_i [\pi^t D_i^t - \lambda_i^t D_i^t]. \quad (8)$$

The parameter $\pi^t D_i^t$ is the grid operator's gross profit, which $\pi^t$ can be different depending on the value of demand reduction for the grid operator in each time slot $t$. Also we define, $\boldsymbol{\lambda}_i = [\lambda_i^1, \lambda_i^2, \ldots, \lambda_i^T]^\top$, $\mathbf{D}_i = [D_i^1, D_i^2, \ldots, D_i^T]^\top \forall i \in \mathcal{N}$ and $\boldsymbol{\lambda} = [\boldsymbol{\lambda}_1^\top, \boldsymbol{\lambda}_2^\top, \ldots, \boldsymbol{\lambda}_N^\top]^\top$, $\mathbf{D} = [\mathbf{D}_1^\top, \mathbf{D}_2^\top, \ldots, \mathbf{D}_N^\top]^\top$.

*Remark 2: The penalty price does not affect the grid operator's utility as it derives no penalty revenue. Specifically, the grid operator does not intend to increase its utility by designing the penalty. The penalty is only a tool to keep customers from deviating from the contracted demand reduction [15].*

### C. Formulating the Contract Between GO and CUs

The grid operator faces the challenge of devising a contract strategy that encompasses demand reduction and the corresponding incentive reward and penalty while lacking complete knowledge about the customers' actual types. These incentives, penalties, and demand reductions in each contract are designed to be a function of the customer's type. The following outlines the procedure for creating contracts:

- **Step 1:** grid operator determines total demand reduction for each time slot based on capacity of its generators and customers' predicted demand.
- **Step 2:** grid operator designs optimal contracts by determining incentive reward, penalty, and demand reduction for each type.
- **Step 3:** customers report their type and receive corresponding contract items $(D_i^t, \lambda_i^t, P_i^t)$.
- **Step 4:** On DR event day, customers perform their preferred demand reduction $\tilde{D}_i^t$.
- **Step 5:** customers are rewarded $\tilde{D}_i^t \lambda_i^t$ and penalized $(D_i^t - \tilde{D}_i^t) P_i^t$.

We use contract theory to model the interaction between the grid operator and C&I customers considering incomplete information [24]. A contract is feasible if the following conditions are met: 1) Incentivize customers to participate in the contract (Individual Rationality). 2) Ensure that customers report their type truthfully (Incentive Compatibility). Following is an accurate definition of these conditions.

*Definition 1 (Individual Rationality (IR)): A contract satisfies the IR constraint if the customer's utility is non-negative by reporting its type truthfully, i.e.*

$$U_{CU,i}^t(D_i^t, \lambda_i^t, P_i^t) \geq 0, \ \forall i \in \mathcal{N}, \forall t \in \mathcal{T}. \tag{9}$$

The IR constraint ensures that the customer accepts to sign the contract. If a rational customer can attain a non-negative utility by selecting the contract that corresponds to their type, they have no reason not to participate in the mechanism [30].

*Definition 2 (Incentive Compatibility (IC)): A contract is IC if the customer reaches equal or higher utility through reporting its type truthfully, i.e.* $\forall t \in \mathcal{T}, \forall i, j \in \mathcal{N}$

$$U_{CU,i}^t(D_i^t, \lambda_i^t, P_i^t) \geq U_{CU,i}^t(D_j^t, \lambda_j^t, P_i^t) \, i \neq j. \tag{10}$$

The left-hand side of the inequality represents the utility of the type-$i$ customer upon selecting type-$i$ contract items, while the right-hand side indicates the utility of the type-$i$ customer upon selecting type-$j$ contract items, where $i \neq j$. The IC constraint ensures that the utility of type-$i$ customer is maximized by choosing the contract that is designed for its actual type, i.e. $D_i^t, \lambda_i^t, P_i^t$. Therefore, a customer will have no incentive to misreport its actual type under IC constraint [30].

In the proposed mechanism, the grid operator maximizes its utility, subject to IR and IC constraints and also the constraints imposed on the system model. Based on the grid operator's demand deficiency, to ensure the feasibility of the designed contracts, $K_r^t \geq \frac{D_{req}^t}{MN}$ and $K_s \geq \sum_{t\in\mathcal{T}} K_r^t$ must satisfy in (1) and (2). Therefore, the optimal mechanism can be obtained by solving the following optimization problem:

$$\begin{aligned}
&\max_{(D_i^t, \lambda_i^t, P_i^t)} U_{GO} = \sum_{i=1}^{N} \sum_{t=1}^{T} M f_i [\pi^t D_i^t - \lambda_i^t D_i^t], \\
&s.t. \quad U_{CU,i}^t(D_i^t, \lambda_i^t, P_i^t) \geq 0 \ \ \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \\
&\qquad U_{CU,i}^t(D_i^t, \lambda_i^t, P_i^t) \geq U_{CU,i}^t(D_j^t, \lambda_j^t, P_i^t) \, \forall i, j \in \mathcal{N}, \\
&\qquad i \neq j, \forall t \in \mathcal{T} \\
&\qquad \sum_{i=1}^{N} M f_i D_i^t = D_{req}^t \ \ \forall t \in \mathcal{T} \\
&\qquad 0 < D_i^t \leq K_r^t \sum_{t\in\mathcal{T}} D_i^t \leq K_s \ \ \forall i \in \mathcal{N}, \forall t \in \mathcal{T}.
\end{aligned} \tag{11}$$

*Remark 3: The results of the proposed contract problem (11) indicate that having precise knowledge of the distribution function of customer types is not essential for fulfilling the IC and IR constraints. However, it is important to note that the utility of the grid operator may be impacted.* Drawing upon Theorem 1, we obtain the optimal penalty based on $D^t$ and $\lambda^t$. By substituting $P_i^t = P_i^{t^\star}$ and $\tilde{D}_i^t = D_i^t$ in the optimization problem (11), we obtain:

$$\max_{\{(D_i^t, \lambda_i^t)\}} U_{GO} = \sum_{i=1}^{N} \sum_{t=1}^{T} M f_i [\pi^t D_i^t - \lambda_i^t D_i^t], \tag{12a}$$

$$s.t. \quad U_{CU,i}^t(D_i^t, \lambda_i^t) \geq 0 \ \ \forall i \in \mathcal{N}, \forall t \in \mathcal{T}, \tag{12b}$$

$$U_{CU,i}^t(D_i^t, \lambda_i^t) \geq U_{CU,i}^t(D_j^t, \lambda_j^t) \, \forall i, j \in \mathcal{N}, \ i \neq j, \forall t \in \mathcal{T}, \tag{12c}$$

$$\sum_{i=1}^{N} M f_i D_i^t = D_{req}^t \ \ \forall t \in \mathcal{T}, \tag{12d}$$

$$0 < D_i^t \leq K_r^t \, \forall i \in \mathcal{N}, \forall t \in \mathcal{T}, \tag{12e}$$

$$\sum_{t\in\mathcal{T}} D_i^t \leq K_s \ \ \forall i \in \mathcal{N}, \forall t \in \mathcal{T}. \tag{12f}$$

## IV. Contract Problem Reformulation

Since the optimization (11) has $N \times T$ IR constraints and $N \times (N-1) \times T$ IC constraints are non-convex, the optimization (12) cannot be solved, straightforwardly. Hence, we simplify the constraints in the following.

*Lemma 1: Under the IC constraints, if* $\alpha_1 < \alpha_2 < \cdots < \alpha_N$*, the IR constraints in (12a) can be reduced to*:

$$U_{CU,N}^t(D_N^t, \lambda_N^t) = 0 \, \forall t \in \mathcal{T}. \tag{13}$$

*Proof:* We define the set $S$ as the set of feasible contracts.

$$S = \{(D_i^t, \lambda_i^t) \mid (12b), (12c), (12d), (12e), (12f)\}$$

From IC constraint we have $U_{CU,i}^t(D_i^t, \lambda_i^t) \geq U_{CU,i}^t(D_N^t, \lambda_N^t)$, specifically:

$$U_{CU,i}^t(D_i^t, \lambda_i^t) \geq D_N^t \lambda_N^t - \alpha_i.(D_N^t)^2.$$

Using $\alpha_i \leq \alpha_N$, $\forall i \in \mathcal{N}$ we can write

$$U_{CU,i}^t(D_i^t, \lambda_i^t) \geq D_N^t \lambda_N^t - \alpha_N.(D_N^t)^2.$$

Thus, we can conclude that $U_{CU,i}^t(D_i^t, \lambda_i^t) \geq U_{CU,N}^t(D_N^t, \lambda_N^t) \geq 0$. This is necessary and sufficient condition for establishing the IR constraints. Using necessary and sufficient conditions we prove that in the set $W \subseteq S$, $\arg\max_{(D_j^t, \lambda_j^t)\in W} U_{GO} = \arg\max_{(D_j^t, \lambda_j^t)\in S} U_{GO}$, where

$$\begin{aligned}
W = \{(D_i^t, \lambda_i^t) \ |U_{CU,N}^t(D_N^t, \lambda_N^t) \\
= 0, (12c), (12d), (12e), (12f)\}.
\end{aligned}$$

The optimization objective function is a decreasing function of $\lambda_i^t$, $\forall i \in \{1, \ldots, N\}$. In the optimal contract, we have

$U^t_{CU,N}(D^t_N, \lambda^t_N) = 0$. Otherwise, the value of $\lambda^t_i$ for each type can be reduced by $\epsilon > 0$, which at the same time satisfies all the constraints of (12) and increases the utility of GO. □

From lemma 1, we infer that if IR constraint for the highest customer's type is binding, then the IR constraints for the other's types will automatically hold under the IC constraints. We will find the necessary and sufficient conditions for IC constraints in the following.

*Lemma 2: Under the IC constraints, we have*:

1) *if* $\alpha_i > \alpha_j$, *then* $D^t_i \leq D^t_j$ *(monotonicity)*
2) *if* $\alpha_i > \alpha_j$, *then* $D^t_i\lambda^t_i \leq D^t_j\lambda^t_j$ $\quad \forall i, j \in \mathcal{N}, \ \forall t \in \mathcal{T}$

*Proof:* According to the IC constraint for CU with $type-i$ and CU with $type-j$, we can obtain:

$$D^t_i\lambda^t_i - \alpha_i.(D^t_i)^2 \geq D^t_j\lambda^t_j - \alpha_i.(D^t_j)^2, \tag{14}$$
$$D^t_j\lambda^t_j - \alpha_j.(D^t_j)^2 \geq D^t_i\lambda^t_i - \alpha_j.(D^t_i)^2. \tag{15}$$

Adding (14) and (15), we have:

$$(\alpha_j - \alpha_i).((D^t_i)^2 - (D^t_j)^2) \geq 0. \tag{16}$$

Thus, if $\alpha_i > \alpha_j$, then $(D^t_i)^2 \leq (D^t_j)^2$, and since $D^t \geq 0$ in (1), we have $D^t_i \leq D^t_j$, which means that monotonicity has been proven. To prove the second part, from (15), we have:

$$\alpha_j.((D^t_i)^2 - (D^t_j)^2) \geq D^t_i\lambda_i - D^t_j\lambda^t_j. \tag{17}$$

If $\alpha_i > \alpha_j$, due to the monotonicity we know that $(D^t_j)^2 \geq (D^t_i)^2$. Applying it to (17), we have $D^t_i\lambda^t_i \leq D^t_j\lambda^t_j$. □

From Lemma 2, a customer with a higher type is less inclined to reduce demand, which makes sense according to the customer's utility function in (3). Furthermore, Lemma 2 implies that the profit it receives from the grid operator for reducing demand should be less for larger values of types. Otherwise, each customer tends to choose a higher type to incur lower cost from demand reduction.

*Theorem 2: If the monotonicity condition is met, then in the optimal contract, for* $\forall t \in \mathcal{T}$ *and* $\forall i \in \{1, \ldots, N-1\}$, *the IC constraint in (11) can be reduced to*

$$U^t_{CU,i}(D^t_i, \lambda^t_i) = U^t_{CU,i}(D^t_{i+1}\,\lambda^t_{i+1}). \tag{18}$$

*Proof:* The IC constraint between type-$i$ and type-$(i+1)$ is:

$$U^t_{CU,i}(D^t_i, \lambda^t_i) \geq U^t_{CU,i}(D^t_{i+1}, \lambda^t_{i+1}), \tag{19}$$

and the IC constraint between type-$i$ and type-$(i-1)$ is:

$$U^t_{CU,i}(D^t_i, \lambda^t_i) \geq U^t_{CU,i}(D^t_{i-1}, \lambda^t_{i-1}). \tag{20}$$

First, we show that with the monotonicity conditions, the IC constraints can be reduced to (19) and (20). Given $\alpha_i < \alpha_{i+1} < \alpha_{i+2}$ and $D^t_i \geq D^t_{i+1} \geq D^t_{i+2}$ from Lemma 2, we have:

$$\alpha_{i+1}.(D^t_{i+1})^2 - \alpha_i.(D^t_{i+1})^2 \geq \alpha_{i+1}.(D^t_{i+2})^2 - \alpha_i.(D^t_{i+2})^2. \tag{21}$$

From the (19), we have

$$\lambda^t_{i+1}D^t_{i+1} - \alpha_{i+1}.(D^t_{i+1})^2 \geq \lambda^t_{i+2}D^t_{i+2} - \alpha_{i+1}.(D^t_{i+2})^2. \tag{22}$$

By combination (21) and (22) we have

$$\lambda^t_{i+1}D^t_{i+1} - \alpha_i.(D^t_{i+1})^2 \geq \lambda^t_{i+2}D^t_{i+2} - \alpha_i.(D^t_{i+2})^2,$$

which is equivalent to

$$U^t_{CU,i}(D^t_{i+1}, \lambda^t_{i+1}) \geq U^t_{CU,i}(D^t_{i+2}, \lambda^t_{i+2}). \tag{23}$$

The combination of (23) and (19), gives the result that $U^t_{CU,i}(D^t_i, \lambda^t_i) \geq U^t_{CU,i}(D^t_{i+2}, \lambda^t_{i+2})$. In this way, we can further prove $U^t_{CU,i}(D^t_i, \lambda^t_i) \geq U^t_{CU,i}(D^t_j, \lambda^t_j), \ \forall j > i$. We can prove $U^t_{CU,i}(D^t_i, \lambda^t_i) \geq U^t_{CU,i}(D^t_j, \lambda^t_j), \ \forall j < i$ in the same way with the help of Equation (20). These results mean that (19) and (20) are sufficient conditions for the IC constraints.

Next, we show that in the optimal contract, (19) and (20) can also be simplified to

$$U^t_{CU,i}(D^t_i, \lambda^t_i) = U^t_{CU,i}(D^t_{i+1}, \lambda^t_{i+1}). \tag{24}$$

To show the optimality, using (19), (20), and monotonicity condition, we prove that in the set $Z \subseteq W$, $\arg\max_{(D^t_j,\lambda^t_j)\in Z} U_{GO} = \arg\max_{(D^t_j,\lambda^t_j)\in S} U_{GO}$, where

$$\begin{aligned} Z = \{(D^t_i, \lambda^t_i) \mid\ & U^t_{CU,N}(D^t_N, \lambda^t_N) = 0, \\ & U^t_{CU,i}(D^t_i, \lambda^t_i) = U^t_{CU,i}(D^t_{i+1}, \lambda^t_{i+1}), \\ & D^t_N \leq D^t_{N-1} \leq \ldots \leq D^t_1, (12d), (12e), (12f).\} \end{aligned}$$

We rewrite (24) as follows:

$$\lambda^t_i D^t_i - \alpha_i.(D^t_i)^2 = \lambda^t_{i+1}D^t_{i+1} - \alpha_i.(D^t_{i+1})^2. \tag{25}$$

From (25), $\alpha_{i+1} > \alpha_i$ and $D^t_i > D^t_{i+1}$, we obtain:

$$\lambda^t_i D^t_i - \lambda^t_{i+1}D^t_{i+1} \leq \alpha_{i+1}.((D^t_i)^2 - (D^t_{i+1})^2). \tag{26}$$

Rearranging the term in (26), we obtain:

$$\lambda^t_i D^t_i - \alpha_{i+1}.(D^t_i)^2 \leq \lambda^t_{i+1}D^t_{i+1} - \alpha_{i+1}.(D^t_{i+1})^2. \tag{27}$$

By comparing (20) and (27), we ensure that equation (20) holds. Also, we rearrange Equation (19) as follows:

$$\lambda^t_i \geq \frac{\alpha_i.(D^t_i)^2 + \lambda^t_{i+1}D^t_i - \alpha_i.(D^t_{i+1})^2}{D^t_i}. \tag{28}$$

The optimization objective function is a decreasing function of $\lambda^t_i$, $\forall i \in \{1, \ldots, N\}$. Based on Lemma 1, we find that in the optimal solution, the CU's utility with higher type is equal to zero, i.e. $D^t_N\lambda^t_N - \alpha_N.(D^t_N)^2 = 0$. Thus, the GO sets $\lambda^t_N = \alpha_N.D^t_N$. Given $\lambda^t_N$, for $i = N-1$ in (28), the GO reduces the value of incentive rate $\lambda^t_{N-1}$ to such an extent that (28) holds in the equal sign. Regarding $\lambda^t_{N-1}$, the GO reduces the value of incentive rate $\lambda^t_{N-2}$ to take the equal sign of (28) for $i = N-2$. This process continues sequentially until $i = 1$, and for each $i$, the GO decreases $\lambda^t_i$ such that the equation (28) binds. Therefore, equation (19) for $i = \{1, \ldots, N-1\}$ are binding and equivalent to equation (24). □

Based on Theorem 2, $N \times (N-1) \times T$ IC inequality constraints are reduced to $(N-1) \times T$ equalities. It means that in the optimal contract, for each type-$i$ customer, it is indifferent to select her actual type or the next type-$(i+1)$. Thus, based on the results of Lemma 1 and Theorem 2, the optimization problem (12a) can be rewritten as follows:

$$\max_{(D^t_i,\lambda^t_i)} U_{GO} = \sum_{i=1}^{N}\sum_{t=1}^{T} Mf_i[\pi^t D^t_i - \lambda^t_i D^t_i],$$

$$\begin{aligned} s.t.\quad & U^t_{CU,N}(D^t_N,\lambda^t_N)=0 \;\; D^t_N \le D^t_{N-1} \le \ldots \le D^t_1 \\ & U^t_{CU,i}(D^t_i,\lambda^t_i)=U^t_{CU,i}(D^t_{i+1},\lambda^t_{i+1}), \\ & \forall i \in \{1,\ldots,N-1\} \\ & \sum_{i=1}^{N} M f_i D^t_i = D^t_{req} \\ & 0 < D^t_i \le K^t_r \; \sum_{t\in\mathcal{T}} D^t_i \le K_s \;\; \forall i \in \mathcal{N}. \end{aligned} \tag{29}$$

According to the first and second constraints of (29), the incentive reward $\lambda^t_i$ can be reformulated as a function of the demand reduction $D^t_i$, as:

$$\lambda^t_i = \frac{\alpha_N.(D^t_N)^2 + \sum_i^N w^t_i}{D^t_i} \; \forall i \in \{1,\ldots,N\}, \forall t \in \mathcal{T}, \tag{30}$$

where

$$w^t_i = \begin{cases} \alpha_i.((D^t_i)^2 - (D^t_{i+1})^2), & \forall i \in \{1,\ldots,N-1\}, \\ 0, & i = N. \end{cases}$$

By substituting (30) in the grid operator's utility function, we can eliminate the dependency of $U_{GO}(\mathbf{D},\boldsymbol{\lambda})$ on $\boldsymbol{\lambda}$ and write $U_{GO}(\mathbf{D},\boldsymbol{\lambda}^\star)$ only in terms of demand reductions, where $\boldsymbol{\lambda}^\star$ is the optimal value of $\boldsymbol{\lambda}$. i.e.

$$\begin{aligned} U_{GO}(\mathbf{D},\boldsymbol{\lambda}^\star) &= \sum_{i=1}^{N}\sum_{t=1}^{T} M f_i[\pi^t D^t_i - \lambda^t_i D^t_i] \\ &= \sum_{i=1}^{N}\sum_{t=1}^{T} M f_i[\pi^t D^t_i - \alpha_N.(D^t_N)^2 - \sum_i^N w^t_i] \\ &= \sum_{i=1}^{N}\sum_{t=1}^{T} M[f_i(\pi^t D^t_i) - R_i(D^t_i)^2], \end{aligned} \tag{31}$$

where

$$R_i = \begin{cases} \alpha_i \sum_{j=1}^{i} f_j - \alpha_{i-1}\sum_{j=1}^{i-1} f_j & \forall i \in \{2,\ldots,N\}, \\ \alpha_1 f_1 & \text{for } i=1\ . \end{cases}$$

Thus, the optimization problem (29) becomes:

$$\max_{(D^t_i)} U_{GO} = \sum_{i=1}^{N}\sum_{t=1}^{T} M[f_i(\pi^t D^t_i) - R_i(D^t_i)^2], \tag{32a}$$

$$s.t.\quad D^t_N \le D^t_{N-1} \le \ldots \le D^t_1 \; \forall t \in \mathcal{T}, \tag{32b}$$

$$\sum_{i=1}^{N} M f_i D^t_i = D^t_{req} \; \forall t \in \mathcal{T} \tag{32c}$$

$$0 < D^t_i \le K^t_r \;\; \forall i \in \mathcal{N} \, \forall t \in \mathcal{T}, \tag{32d}$$

$$\sum_{t\in\mathcal{T}} D^t_i \le K_s \;\; \forall i \in \mathcal{N}. \tag{32e}$$

The objective function of optimization (32a) is strictly concave, and also all constraints (32b), (32c), (32d), and (32e) are convex. Therefore the optimization (32) is strictly convex.

**Algorithm 1** Dual Decomposition Algorithm

**Result: D**
**Initialize:** $k \leftarrow 1$, $\mathbf{h}_1 \leftarrow \mathbf{h}_{init}$, $\hat{\mathbf{D}}_1 \leftarrow \hat{\mathbf{D}}_{init}$
**repeat**
1. Solve the dual subproblems
   for each $i \in \mathcal{N}$ and $t \in \mathcal{T}$ solve quadratic subproblem $g_{i,t}(\mathbf{h}_k)$ in (34) and calculate $\hat{D}^t_i$ using the gradient method.
2. $\hat{\mathbf{D}}_{k+1} \leftarrow \hat{\mathbf{D}}$
3. $\eta_k \leftarrow \frac{1}{k}$
4. Updates lagrangian coefficients
   $\mathbf{h}_{k+1} = (\mathbf{h}_k + \eta_k \mathbf{b}(\hat{\mathbf{D}}_{k+1}))_+$
5. $k \leftarrow k+1$

**until** $\|\hat{\mathbf{D}}_k - \hat{\mathbf{D}}_{k-1}\| \le \epsilon_{stop}$;

## V. Solution of Optimal Contract

Due to the existence of coupling constraints (32b), (32c) and (32e) in the optimization problem and to avoid the high computational cost imposed by a centralized method for large-scale systems, we use the dual decomposition method to solve the optimization problem (32) in an iterative fashion.

The Lagrangian function of the optimization (32) is:

$$\begin{aligned} L(\mathbf{D},\mathbf{h}) &= \sum_{i=1}^{N}\sum_{t=1}^{T}(-M[f_i(\pi^t D^t_i) - R_i(D^t_i)^2]) + \mathbf{h}^\top \mathbf{b}(\mathbf{D}) \\ &= \sum_{i=1}^{N}\sum_{t=1}^{T} G(D^t_i,\mathbf{h}), \end{aligned} \tag{33}$$

where $\mathbf{b}$ is the vector of $3NT+N$ linear inequality constraint in (32b), (32c), (32d), (32e), and $\mathbf{h}$ is a vector of Lagrange coefficients corresponding to each constraint in $\mathbf{b}$. The dual sub-problems is defined as follows:

$$g_{i,t}(\mathbf{h}) = \inf_{D^t_i} \; G(D^t_i,\mathbf{h}) \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N}. \tag{34}$$

After the agents solve all sub-problem and find $\hat{\mathbf{D}} = \{\hat{\mathbf{D}}_1, \hat{\mathbf{D}}_2, \ldots, \hat{\mathbf{D}}_N\}$, the coordinator updates the Lagrangian coefficients vector $\mathbf{h}$ based on $\mathbf{b}(\hat{\mathbf{D}})$ as follows:

$$\mathbf{h} := (\mathbf{h} + \eta \mathbf{b}(\hat{D}))_+, \tag{35}$$

where $\eta$ is subgradient step size and $(.)_+ = max(.,0)$. According to Proposition 6.3.1 in [31], it is proved that the Lagrangian coefficients converge to the optimal value if $\eta_k \ge 0$ , $\sum_{k=0}^{\infty} \eta_k = \infty$, $\sum_{k=0}^{\infty} \eta_k^2 \le \infty$ and $\eta_\infty = 0$.

The procedure of the proposed Dual decomposition is presented in algorithm 1. This algorithm obtains demand reductions by solving the optimization (32). Then, using (30), the incentive reward of each demand reduction are calculated.

*Remark 4: The calculation of the optimal demand reduction from Algorithm 1 has a complexity of $O(N^2)$, where N is the number of customers' types.*

## VI. Simulation

The numerical simulation results are presented to evaluate the proposed incentive-based DR scheme's performance. We consider a scenario that includes one grid operator

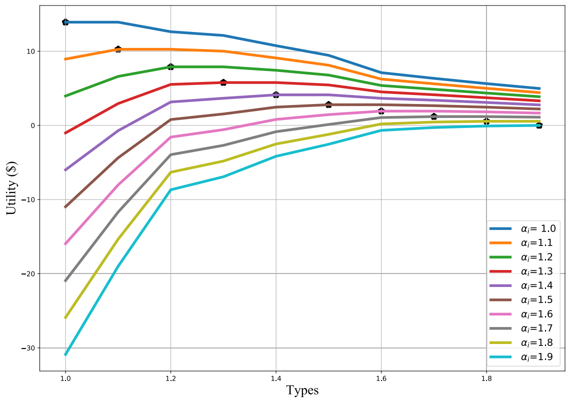



Fig. 2. Utilities of 10 types of customer $i$ when reporting different types as their types.

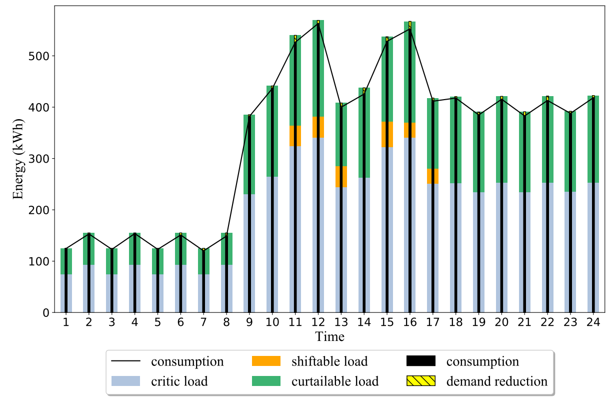
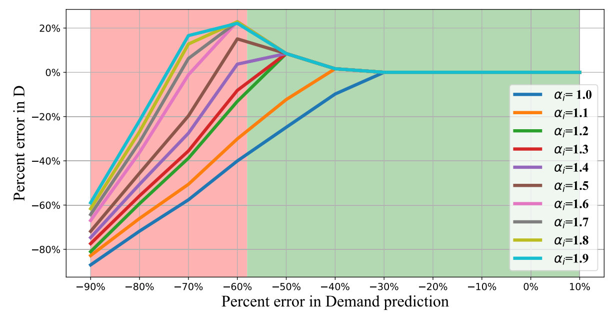
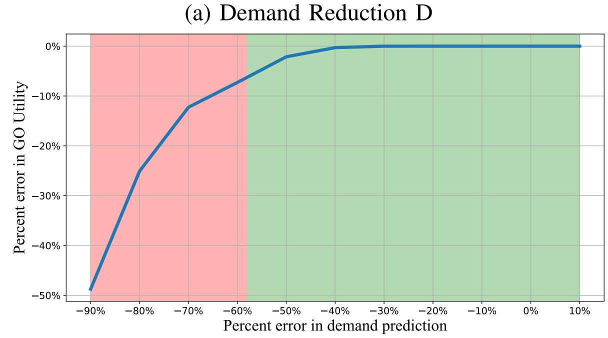



Fig. 3. The optimal customer's demand reduction with type $\alpha_i = 1.4$ in his load profile.




(a) Demand Reduction D




(b) GO Utility

Fig. 4. Analyzing the impact of load prediction errors on contract design results.

and 24 C&I customers. The datasets used for customers are selected from [32]. This dataset contains the hourly load profiles over a year (8760 hours of data) for 24 representative facilities from various end-use sectors. We assume that there are 10 different types of customers in the grid, whose types of customers are chosen from the set $\Delta = \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9\}$ $\$/(kWh)^2$. The $\pi^t$ and $D^t_{req}$ interval of the grid operator were set to [18 - 40] $\$/(kWh)$ and [10 - 440] $kWh$, respectively.

### A. Properties of Proposed Contract

To validate the feasibility of the proposed contract, the utilities of all types of customers, when they reported different types, are depicted in Figure 2. The black stars marked on the curve show the points where each type of customer gains maximum utility. Figure 2 shows that each customer makes the most profit when it reports its actual type. In this way, customers have no incentive to misreport their actual type, indicating that the IC constraint is satisfied. Moreover, when customers choose contract items of their actual type, they gain a non-negative utility. Such results verify the feasibility of the proposed contract under information asymmetry.

In Figure 3, we show the amount of demand reduction that one of the customers accepts during one day of the event days. As one of the customers, 24-hour demand profile of the Industrial-Food Processing Facility with type $\alpha_i = 1.4$ is shown in Figure 3. As shown in Figure 3, the customer's demand reduction is such that it has both the condition of providing the critic load and performing the shiftable load during the day.

Figure 4 illustrates the impact of prediction errors in parameters associated with load types (i.e., $K^t_r$ and $K_s$) on customers' load reduction. Figure 4a shows that as prediction errors increase, customers with smaller type are impacted the most. This is because they can more easily meet demand reduction. Narrower prediction errors lead to more significant demand reduction for them. Additionally, Figure 4b shows that increased prediction errors result in reduced profit for the grid operator. The results remain stable up to a $-30\%$ error, and contracts are feasible up to a $-58\%$ error. However, surpassing this threshold caused the failure of the demand deficiency constraint and made the contracts unfeasible.

In Figure 5, we examine how customers' preferred energy reductions affect contracts with and without penalties. Suppose $\tilde{D}^t_i = e^t D^t_i$, where $e^t$ represents customers' different execution levels from the contracted reduction. Customers' execution levels range from [0, 1], determined by themselves after signing the contract. When $e^t = 0$, they provide no reduction, and when $e^t = 1$, they meet or exceed the contracted reduction. In the scenario without penalties, customers can achieve more profit if they unexpectedly reduce their performance. However, in the penalty scenario, their profit decreases for each deviation from the contract, and they pay a penalty if they can't fully execute. When a customer's execution level falls below $1-\sqrt{(\frac{\lambda^t_i}{\alpha_i D^t_i}-1)}$, they obtain a negative profit. Full execution in both scenarios results in the same profit, indicating an effective penalty design to uphold the contracts.

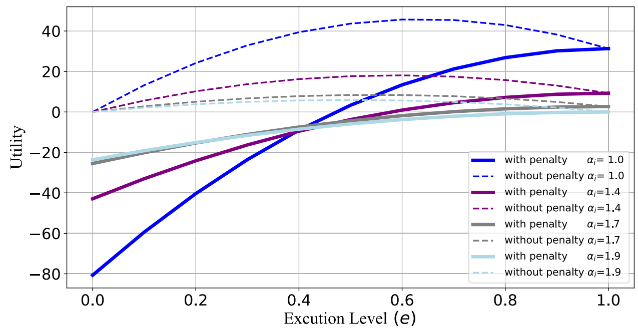
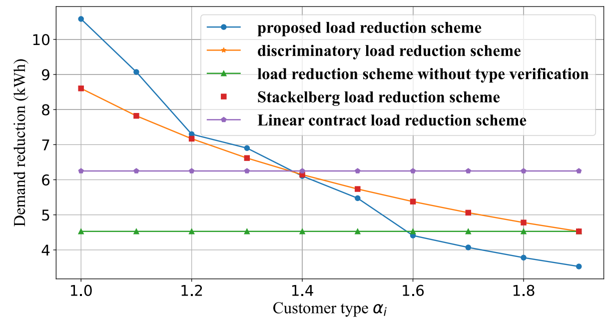



Fig. 5. Utility of customers under different execution level.

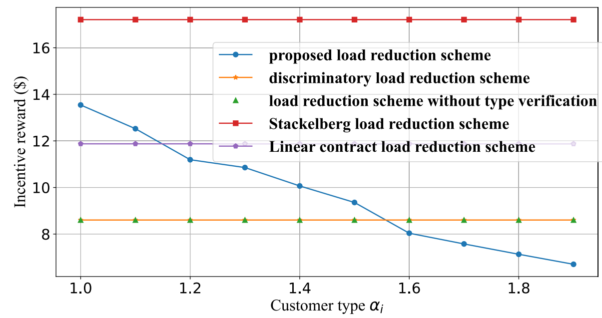
## B. Performance Comparison

To evaluate the benefits of the proposed load reduction scheme, we compare it with the discriminatory load reduction scheme [33], load reduction scheme without type verification [34], Stackelberg load reduction scheme [28], and linear contract load reduction scheme [22]. The discriminatory load reduction scheme is obtained from the optimal contract under complete information, i.e., the grid operator is aware of the types of customers. In the load reduction scheme without type verification, the grid operator does not know the types of customers, and it assumes all customers report their type truthfully. However, the agents can misreport their type to maximize their profit. The Stackelberg load reduction scheme involves the grid operator is informed about the customer types while the customers determine the amount of demand reduction. The grid operator first determines the amount of incentive rewards for each customer, and then the customers have the freedom to curtail their demand as per their preference. The linear contract load reduction scheme considers a linear form in terms of demand reduction for the incentive function and solves the problem with the contract theory approach.

In Figure 6, we illustrate the optimal demand reduction and incentive rewards for all customer types across all load reduction schemes. In Figure 6a, we see that in the proposed scheme, the demand reduction decreases with increasing customer type, confirming the monotonicity constraint in Lemma 2. We also see that the amount of demand reduction in the discriminatory load reduction scheme is equal to that in the Stackelberg scheme because both have the same objective functions, leading to identical optimal outcomes for demand reduction. Both schemes offer greater incentives to customers with lower types to reduce their demand. However, in the Stackelberg scheme, the grid operator provides more incentives than the discriminatory scheme to encourage customers to choose a reduction amount that is more responsive to the expected demand deficit. In the load reduction without type verification scheme, since the IC constraint is not considered, customers of any type can choose the contract that brings the most profit. As a result, the amount of demand reduction and incentive rewards are the same for all types of customers. Under the assumption of linearity of incentive function based on demand reduction, the establishment of IR and IC constraints in linear contract scheme requires $\mu = \alpha_N$, $D_N^t = D_{N-1}^t = \ldots = D_1^t \; \forall t \in \mathcal{T}$, respectively. Thus, demand reduction and incentive reward are equal to a fixed value.




(a) The optimal value of demand reduction vs. Type of customers.




(b) The optimal value of incentive reward vs. Type of customers.

Fig. 6. The optimal contract $(D^\star, \lambda^\star)$ for different types of customers under different schemes in $t = 11$.

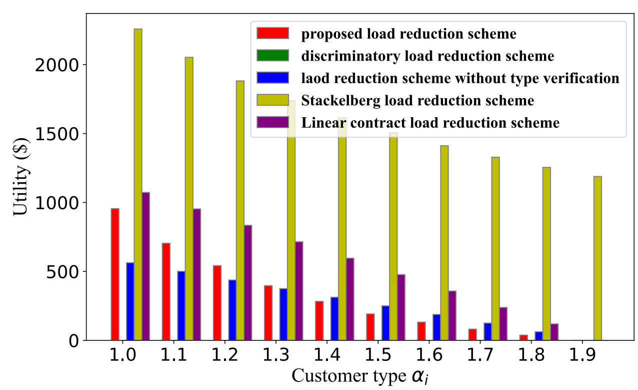
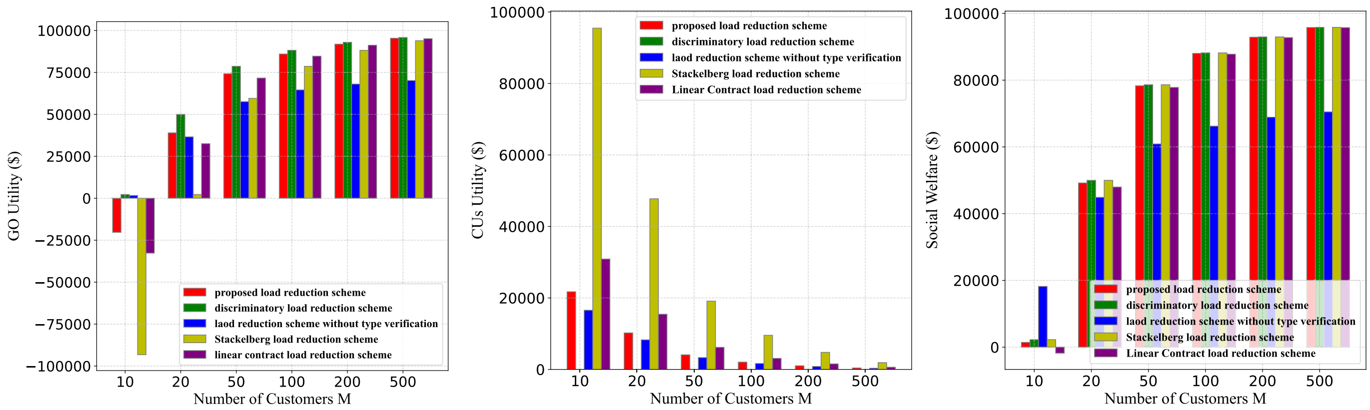


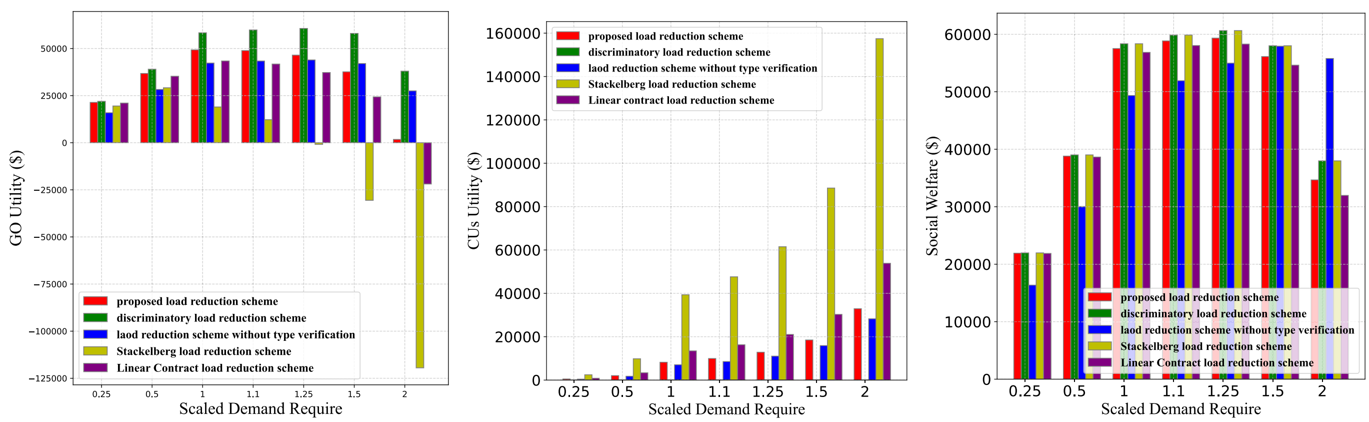

Fig. 7. Utility of customers in different schemes.

Figure 7 investigates the utilities of various customer types across all comparison schemes. In the Stackelberg load reduction scheme, customers independently decide on the amount of demand reduction, maximizing their potential profit. In contrast, the discriminatory load reduction scheme has the grid operator determining each customer type's demand reduction, only satisfying the IR constraints, resulting in zero utility for all customer types. In the load reduction without type verification scheme, the grid operator disregards the IC constraint. Consequently, strategic customers report the highest possible type to increase their profits. Lower-type customers experience less dissatisfaction and greater utility in this scheme. Moreover, we observe that the linear contract scheme outperforms the proposed approach in terms of CUs utility. This is mainly because the linear contract mechanism incorporates a specific structure for rewarding, resulting in greater payments to customers to meet IC and IR constraints.




(a) GO Utility (b) CUs Utility (c) Social Welfare

Fig. 8. System performance with respect to numbers of CUs.




(a) GO Utility (b) CUs Utility (c) Social Welfare

Fig. 9. System performance with respect to various demand deficiency of GO.

Table I compares the grid operator's utility and customers' utility based on the contracts signed during an event day. Under the discriminatory load reduction scheme, customers' total utilities remain zero, resulting in the highest profit for the grid operator. This is because the grid operator has full knowledge of customer types and designs discriminatory contracts, leaving all customers with zero utility to meet the IR constraints. The Stackelberg scheme yields the highest total utilities for customers but the lowest profit for the grid operator. In this scheme, the grid operator must compensate for the expected demand deficit, resulting in more incentive rewards for customers and less profit for itself. The proposed scheme and load reduction without type verification scheme require the grid operator to adjust contracts with incomplete customer information. The grid operator has more profit in the proposed scheme compared to the load reduction without type verification scheme, while the profit of customers in the proposed scheme is less. The utility obtained by the grid operator in our proposed scheme surpasses that of the linear contract scheme. The reason behind this is no limitation on the reward function in the proposed scheme and to underscore the efficacy of our optimal reward design. Compared with the linear contract scheme, our scheme excels at achieving the primary objective of maximizing the contract designer's profit.

TABLE I

UTILITIES OF GRID OPERATOR AND CUSTOMERS UNDER DIFFERENT SCHEMES

| **Load Reduction Schemes** | **GO's utility** | **CUs' utility** |
|---|---|---|
| Proposed scheme | 49297.8 | 8222.4 |
| Discriminatory scheme | 58355.9 | 0 |
| Scheme without type verification | 42279.3 | 8367.3 |
| Stackelberg scheme | 18991.9 | 39364.0 |
| Linear Contract scheme | 43387.9 | 13463.8 |

### *C. Scalability Analysis*

In Fig 8, we evaluate the utilities of the GO, CUs, and overall social welfare as the number of CUs varies. As the number of customers increases, the individual burden of demand reduction decreases, leading to a proportional reduction in GO's reward to each customer and an increase in GO's utility (Figure 8a). At the same time, customers experience a decline in their utilities (Figure 8b). With a small number of customers, GO's utility may even turn negative, requiring strategic allocation of the required demand reduction among a limited number of customers. As customers are compelled to undertake a substantial demand reduction, meeting the IR and IC constraints requires that GO pay a significant reward, incurring losses in the situation. However, as the

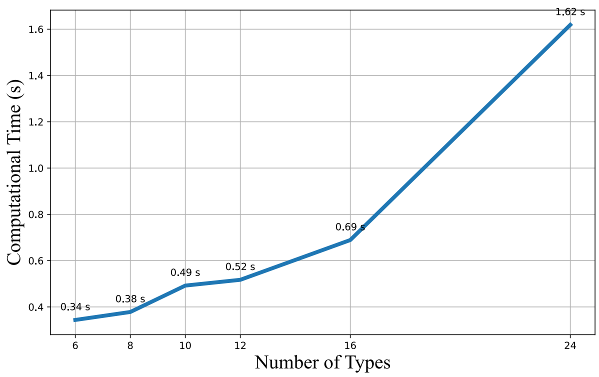



Fig. 10. Time complexity analysis.

number of customers increases, social welfare increases in all schemes (Figure 8c). Our proposed scheme under incomplete information progressively converges towards optimal benchmarks as the number of customers increases. Compared with the linear contract scheme, as depicted in Figures 8c, while linear rewarding may enhance customers' utilities (albeit not obligatory), our scheme outperforms in terms of social welfare.

In Figure 9, we investigate how changing the demand deficit value affects the utility of the GO, CUs, and social welfare. As $w$, the parameter weighting the value of $D_{req}$, increases, the utility of GO initially rises but then decreases (9a). This decline is due to the higher rewards GO must pay for the large volume of demand reduction. Adjusting the valuation of demand reduction ($\pi$) can alter this outcome. As the volume of demand reductions increases, the rewards to customers rise, increasing their utility (Figure 9b). Figure 9c illustrates that with an increase in $w$, the difference between social welfare in the proposed scheme and social welfare in other schemes should decrease. However, a significant increase in $w$ leads to a decrease in social welfare due to a decrease in GO profit.

The impact of the number of customers' type on the calculation time for optimal demand reduction is displayed in Figure 10 which supports our statement in Remark 4. It shows that the time complexity of the problem does not depend on the number of customers ($M$); instead, we categorize the contracts based on the type of customers. Thus, the increase in the number of customers does not affect the computational complexity, but only the amount of items in the contracts. This distinction is crucial, as it highlights the model's resilience to scalability challenges, providing a foundation for its applicability in scenarios with diverse customer bases without compromising computational efficiency.

## VII. Conclusion

In this paper, we proposed an innovative incentive-based Demand Response (DR) mechanism for Commercial and Industrial customers, addressing the challenges arising from incomplete information on customer preferences and the need for effective customer participation. Our approach leveraged the potential of contract theory to design a DR program that accommodated various kinds of customer loads, including curtailable, critical, and shiftable loads while maximizing the grid operator's profit. We overcame the inherent non-convexity of the optimization problem derived from contract theory by employing a mathematical reformulation technique, transforming the problem into a tractable and convex one with a reduced set of constraints. The results showed that the proposed scheme had the best performance in terms of maximizing the grid operator's profit compared to other incomplete information schemes. Meanwhile, it was shown that optimal penalty design could prevent customers from deviating from contracts. Extensive numerical evaluations demonstrated the proposed framework's effectiveness compared to existing benchmark schemes, emphasizing its potential in successfully implementing DR programs for large non-residential customers, consequently reducing the peak load on the grid.

In future work, we will explore the integration of the proposed mechanism with other grid management techniques, such as energy storage and renewable energy sources, to further improve the efficiency and sustainability of the power system.

## References

[1] Z. Wang et al., "How to effectively implement an incentive-based residential electricity demand response policy? Experience from large-scale trials and matching questionnaires," *Energy Policy*, vol. 141, Jun. 2020, Art. no. 111450.

[2] S. M. Azzam, T. Elshabrawy, and M. Ashour, "A bi-level framework for supply and demand side energy management in an islanded microgrid," *IEEE Trans. Ind. Informat.*, vol. 19, no. 1, pp. 220–231, Jan. 2023.

[3] D. Liu, Z. Qin, H. Hua, Y. Ding, and J. Cao, "Incremental incentive mechanism design for diversified consumers in demand response," *Appl. Energy*, vol. 329, Jan. 2023, Art. no. 120240.

[4] P. Siano, "Demand response and smart grids—A survey," *Renew. Sustain. Energy Rev.*, vol. 30, no. 2, pp. 461–478, 2014.

[5] M. Yu, S. H. Hong, and J. Beom Kim, "Incentive-based demand response approach for aggregated demand side participation," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Nov. 2016, pp. 51–56.

[6] *Emergency Load Reduction Program*. Accessed: Mar. 19, 2024. [Online]. Available: https://www.cpuc.ca.gov/industries-and-topics/electrical-energy/electric-costs/demand-response-dr/emergency-load-reduction-program

[7] *Commercial Load Management Program*. Accessed: Mar. 19, 2024. [Online]. Available: https://www.katyedc.org/for-site-selectors/incentives-financing-and-taxes/p/item/1581/centerpoint-energys-commercial-load-management-program

[8] H. Golmohamadi, "Demand-side management in industrial sector: A review of heavy industries," *Renew. Sustain. Energy Rev.*, vol. 156, Jan. 2022, Art. no. 111963.

[9] M. Yu, S. H. Hong, Y. Ding, and X. Ye, "An incentive-based demand response (DR) model considering composited DR resources," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1488–1498, Feb. 2019.

[10] M. Montazeri, H. Kebriaei, and B. N. Araabi, "A tractable truthful profit maximization mechanism design with autonomous agents," *IEEE Trans. Autom. Control*, early access, 2024.

[11] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*. New York, NY, USA: Cambridge Univ. Press, 2007.

[12] A. Falsone, K. Margellos, S. Garatti, and M. Prandini, "Dual decomposition for multi-agent distributed optimization with coupling constraints," *Automatica*, vol. 84, pp. 149–158, Oct. 2017.

[13] R. Besharati, M. H. Rezvani, and M. M. G. Sadeghi, "An incentive-compatible offloading mechanism in fog-cloud environments using second-price sealed-bid auction," *J. Grid Comput.*, vol. 19, no. 3, pp. 1–29, Sep. 2021.

[14] D. An, Q. Yang, D. Li, and Z. Wu, "Distributed online incentive scheme for energy trading in multi-microgrid systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 21, no. 1, pp. 951–964, Jan. 2024.

[15] D. Muthirayan, D. Kalathil, K. Poolla, and P. Varaiya, "Mechanism design for demand response programs," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 61–73, Jan. 2020.

[16] D. Muthirayan, E. Baeyens, P. Chakraborty, K. Poolla, and P. P. Khargonekar, "A minimal incentive-based demand response program with self reported baseline mechanism," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2195–2207, May 2020.

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

[18] P. Samadi, H. Mohsenian-Rad, R. Schober, and V. W. S. Wong, "Advanced demand side management for the future smart grid using mechanism design," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1170–1180, Sep. 2012.

[19] J. Wang, H. Zhong, Z. Yang, X. Lai, Q. Xia, and C. Kang, "Incentive mechanism for clearing energy and reserve markets in multi-area power systems," *IEEE Trans. Sustain. Energy*, vol. 11, no. 4, pp. 2470–2482, Oct. 2020.

[20] G. Tsaousoglou, K. Steriotis, N. Efthymiopoulos, P. Makris, and E. Varvarigos, "Truthful, practical and privacy-aware demand response in the smart grid via a distributed and optimal mechanism," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3119–3130, Jul. 2020.

[21] N. Irtija, F. Sangoleye, and E. E. Tsiropoulou, "Contract-theoretic demand response management in smart grid systems," *IEEE Access*, vol. 8, pp. 184976–184987, 2020.

[22] N. Patrizi, S. K. LaTouf, E. E. Tsiropoulou, and S. Papavassiliou, "Prosumer-centric self-sustained smart grid systems," *IEEE Syst. J.*, vol. 16, no. 4, pp. 6042–6053, Dec. 2022.

[23] L. Ma, N. Liu, H. Sun, C. Li, and W. Liu, "Bi-level frequency regulation resource trading for electricity consumers: A data-driven contract approach," *Int. J. Electr. Power Energy Syst.*, vol. 135, Feb. 2022, Art. no. 107543.

[24] B. Zhang, C. Jiang, J.-L. Yu, and Z. Han, "A contract game for direct energy trading in smart grid," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2873–2884, Jul. 2018.

[25] U. Amin, M. J. Hossain, W. Tushar, and K. Mahmud, "Energy trading in local electricity market with renewables—A contract theoretic approach," *IEEE Trans. Ind. Informat.*, vol. 17, no. 6, pp. 3717–3730, Jun. 2021.

[26] J. Zhu and Y. Zhang, "How to balance the industrial customers' resources requirements while maintaining energy efficiency?" *J. Innov. Knowl.*, vol. 8, no. 1, Jan. 2023, Art. no. 100301.

[27] H. Farzaneh, M. Shokri, H. Kebriaei, and F. Aminifar, "Robust energy management of residential nanogrids via decentralized mean field control," *IEEE Trans. Sustain. Energy*, vol. 11, no. 3, pp. 1995–2002, Jul. 2020.

[28] M. Yu and S. H. Hong, "Incentive-based demand response considering hierarchical electricity market: A Stackelberg game approach," *Appl. Energy*, vol. 203, pp. 267–279, Oct. 2017.

[29] Z. Xiong, J. Zhao, Y. Zhang, D. Niyato, and J. Zhang, "Contract design in hierarchical game for sponsored content service market," *IEEE Trans. Mobile Comput.*, vol. 20, no. 9, pp. 2763–2778, Sep. 2021.

[30] O. D. Hart and B. Holmstrm, "The theory of contracts," Dept. Econ., Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep., 1986.

[31] D. P. Bertsekas, "Nonlinear programming," *J. Oper. Res. Soc.*, vol. 48, no. 3, p. 334, Mar. 1997.

[32] F. Angizeh, A. Ghofrani, and M. Jafari, "Dataset on hourly load profiles for a set of 24 facilities from industrial commercial and residential end-use sectors," *Mendeley Data*, vol. 1, Jan. 2020.

[33] A. Jadbabaie and A. Kakhbod, "Optimal contracting in networks," *J. Econ. Theory*, vol. 183, pp. 1094–1153, Sep. 2019.

[34] M. Montazeri, H. Kebriaei, B. N. Araabi, and D. Niyato, "Optimal mechanism design in the sponsored content service market," *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 3051–3054, Sep. 2021.

**Sajad Parvizi** received the B.S. degree in electrical engineering from Shiraz University of Technology, Shiraz, Iran, in 2019, and the M.S. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 2023. His research interests include game theory, mechanism design, machine learning, and optimization.

**Mina Montazeri** received the Ph.D. degree in control systems from the University of Tehran, Iran, in 2023. She is currently a Post-Doctoral Researcher with the Urban Energy Systems Laboratory, Swiss Federal Laboratories for Materials Science and Technology (Empa), Dübendorf, Switzerland. Her research interests include reinforcement learning, mechanism design, reinforcement learning, optimization, and smart grids.

**Hamed Kebriaei** (Senior Member, IEEE) received the Ph.D. degree in control systems from the University of Tehran, Iran, in 2010. He is currently an Associate Professor of control systems with the School of Electrical and Computer Engineering, University of Tehran. His research interests include game theory, distributed optimization, and multi agent reinforcement learning. He is a TC Member of IEEE CSS in Networks and Communication Systems and a Board Member of the Control Systems Chapter, IEEE Iran Section. He received the Outstanding Reviewer Award from IEEE TRANSACTIONS ON CYBERNETICS in 2022. He received the Outstanding Young Researcher Award from the University of Tehran in 2023. He has served as the Guest Editor for IEEE CONTROL SYSTEMS LETTERS in 2023.