Operator-as-a-Consumer: A Novel Energy Storage Sharing Approach Under Demand Charge

Bingyun Li, Qinmin Yang, Lingjie Duan, and Youxian Sun.

Abstract—Energy Storage Systems (ESS) based Demand Response (DR) is an appealing way to save electricity bills for consumers under demand charge and Time-of-Use (TOU) price. In order to counteract the high investment cost of ESS, a novel operator-enabled ESS sharing scheme, namely, the “Operator-as-a-Consumer (OaaC)”, is proposed and investigated in this paper. In this scheme, the users and the operator form a Stackelberg game. The users send ESS orders to the operator and apply their own ESS dispatching strategies for their own purposes. Meanwhile, the operator maximizes its profit through optimal ESS sizing and scheduling, as well as pricing for the users’ ESS orders. The feasibility and economic performance of OaaC are further analyzed by solving a bi-level joint optimization problem of ESS pricing, sizing, and scheduling. To make the analysis tractable, the bi-level model is first transformed into its single-level Mathematical Program with Equilibrium Constraints (MPEC) formulation, and is then linearized into a Mixed-Integer Linear Programming (MILP) problem using multiple linearization methods. Case studies with actual data are utilized to demonstrate the profitability for the operator and simultaneously the ability of bill saving for the users under the proposed OaaC scheme.

Index Terms—Demand charge, Energy Storage Systems (ESS) sharing, Stackelberg game, bi-level optimization, mathematical program with equilibrium constraints (MPEC), reformulated normalized multiparametric disaggregation (RNMDT).

NOMENCLATURE

Indices:

- $i \in \{1, ..., N\}$ Index of the users
- $j \in \{1, ..., J\}$ Index of dual variables
- $l \in \{0, ..., Q\}$ Index for binary digits in linearization
- $m \in \{0, ..., D\}$ Index of days
- $t \in \{0, ..., T\}$ Index of time slots

Variables:

- $c_{p,cap}$ Unit price for ESS power capacity orders
- $c_{e,cap}$ Unit price for ESS energy capacity orders
- $e^i$ Stored energy of user $i$
- $E$ Stored energy of the operator
- $E_b$ "Useful” stored energy of the operator
- $E_{cap}$ ESS energy capacity ordered by user $i$
- $E_{cap}^i$ ESS energy capacity of the operator $i$
- $g^i$ Dual objective of user $i$
- $o^i$ Primal objective of user $i$
- $O$ Objective of the operator
- $p_{cap}^i, p_d^i$ Charging/discharging power orders of user $i$
- $p_{cap}$ ESS power capacity ordered by user $i$
- $\hat{p}_{cap, k, l}^i$ Linear estimation of $p_{cap}^i p_{cap, k, l}^i$
- $p_l^i$ "Useful” battery power of the operator
- $P_b$ ESS inverter power capacity of the operator
- $P_{max}$ Peak power demand of the operator
- $P_s$ A slack variable
- $w_i^p$ Linear estimation of $c_{p,cap} p_{cap}^i$
- $w_i^e$ Linear estimation of $c_{e,cap} e_{cap}^i$
- $Y$ Concatenated vector of decision variables of the linearized single-level problem
- $Y_l^i$ Concatenated vector of decision variables of user $i$’s lower-level problem
- $Y_u$ Concatenated vector of decision variables of the upper-level problem
- $z_{p, l}$ Binary selectors in RNMDT
- $\lambda^i_p, \mu^i_p$ Dual variables
- $\lambda$ An auxiliary variable in RNMDT
- $\Delta \lambda_p, \Delta u_p^i$ Cut-off errors

Parameters:

- $c_{p,cap, L}, c_{p,cap, H}$ Lower/upper bounds of $c_{p,cap}$
- $C_{P,cap}$ Unit price of physical ESS power capacity
- $C_{E,cap}$ Unit price of physical ESS energy capacity
- $D$ Number of days in the optimization horizon
- $f^i$ Minimum cost of user $i$ using individual ESS
- $k, k_d$ Charge/discharge rates of ESS battery
- $N$ Number of the users
- $p_{cap, L}, p_{cap, H}$ Lower/upper bounds of $p_{cap}^i$
- $p_L^i$ Load profile of user $i$
- $Q$ Cut-off digit
- $r$ TOU price $$/kW$
- $r_m$ Demand charge rate $$/kW$
- $r_{off,peak}$ Off-peak price in TOU price
- $r_{peak}$ Peak price in TOU price
- $soc^i_{ini}$ Initial SOC of user $i$
- $soc_{max}$ Maximum allowed SOC of user $i$
- $soc_{min}$ Minimum allowed SOC of user $i$
- $SOC^i_{ini}$ Initial SOC of the operator
- $SOC_{max}$ Maximum allowed SOC of the operator

This work is supported by the Zhejiang Provincial Nature Science Foundation of China under Grant LZ21F030004, Key-Area Research and Development Program of Guangdong Province under Grant No. 2018B010107002 (Corresponding author: Qinmin Yang)

Bingyun Li, Qinmin Yang, and Youxian Sun are with the State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University, Hangzhou, Zhejiang, 310027 China (e-mail: qnyang@zju.edu.cn).

Lingjie Duan is with the Engineering Systems and Design Pillar at Singapore University of Technology and Design, Singapore (e-mail: lingjie_duan@sutd.edu.sg).
**I. INTRODUCTION**

Demand charge and Time-of-Use (TOU) price are common electricity rate schemes applied to various electricity consumers [1]. The demand charge scheme charges consumers for their maximum grid-side power demand (kW) in each month, while TOU price sets different prices for energy consumption (kWh) at different hours. Hence, consumers are prone to shift some demands to off-peak hours as well as reduce their peak power, namely demand response (DR) [2]–[6]. Using behind-the-meter Energy Storage System (ESS) can significantly enhance DR performance [7]–[13]. However, the high price of ESS casts a shadow on the feasibility of ESS-based DR, or even makes it unprofitable [14]. Inspired by the concept of Sharing Economy, this obstacle leads to the development of ESS sharing. The basis of ESS sharing schemes lies in the flexibility brought by smart grid infrastructures [15]–[19]: The grid company can charge its electricity consumers according to their “modified” load profiles instead of their measured loads.

Various studies on ESS sharing have been proposed in the literature. One possible way is a group of users sharing their individual ESSs with other group members. Recently, this interesting idea has been further extended to the concept of “energy sharing”, where peer-to-peer (p2p) energy trading is enabled among a cluster of prosumers [20]. Based on the inherent distributed structure of p2p energy sharing, and autonomy and selfishness of participants, game-theoretical approaches [21], [22] become the most common ways for developing p2p energy sharing schemes. Other possible approaches include distributed optimization methods [23] and other optimization approaches [20]. Generally, p2p sharing is flexible since users are fully interconnected via information links. However, problems such as power routing and complex market mechanisms cast a shadow on the implementation of p2p energy sharing. Hence, current projects are still mostly confined to the wholesale market [24].

Another promising approach is a group of users sharing a centralized community-scale ESS together. This approach is simple in structure and thus more practical for community-scale applications. Fleischhacker et al. [25] considered the scenario where the ESS is owned by an operator. Through a bi-level optimization approach, it was verified that both the operator’s profit and the users’ cost savings can be achieved. Chakraborty et al. [26] studied the cases where the users share their own ESS or deploy a large-scale ESS together, using cooperative game theory. It is concluded that cooperation is beneficial for the users in both cases.

Liu et al. [14], [15], [27], [28] developed the Cloud Energy Storage (CES) mode, where the concept of “virtual ESS” was involved to enhance the flexibility and improve the social welfare. However, in these studies, the users’ autonomy is neglected, i.e., the users have to exactly follow the solutions from the operator or an aggregator, which may cause conflicts with their own desires. Moreover, the users’ privacy is not preserved by releasing their real-time load information to the aggregator/operator. To address such issues, Oh et al. [29] discussed an optimal ESS dividing strategy for the users’ cost minimization. The ESS was divided into parts, each one of which can be scheduled by a specific user. Yang et al. [9] formulated a joint scheduling problem of ESS and elastic loads. The users can set their own objective functions as well as select their involved elastic loads. However, the ESS deploying costs and demand charge were not considered and cannot be integrated into these schemes easily.

A novel ESS sharing scheme is proposed and analyzed in this paper to deal with these issues. The contributions of this paper are summarized as follows.

a) A novel Stackelberg-game-based ESS sharing scheme, named “Operator-as-a-Consumer (OaaC)” operating mode, is proposed for multiple electricity consumers. Compared to existing operating modes, this mode considers all participants’ autonomy and privacy concerns since it is compatible with users’ own ESS dispatching methods. Meanwhile, it considers the demand charge tariff which brings coupling effect.

b) Offline analyses of the OaaC mode are conducted. The feasibility and economic performance of the OaaC mode are demonstrated by formulating and analyzing the Stackelberg game in a bi-level OaaC operator profit maximization model.

c) A novel model linearization technique is proposed to deal with nonlinear constraints without introducing any integer variables. With the help of this technique and other methods such as the Reformulated Normalized Multiparametric Disaggregation (RNMDT) method, a single-level MILP model with fewer integers is innovatively derived.

The MILP model is finally solved using CPLEX. Benefits in both the users’ bill savings and the operator’s profitability are shown by multiple case studies, using actual load data and electricity prices in practice. The effects of linearization accuracy are also carefully examined.

The rest of this paper is organized as follows. The OaaC mode is introduced in detail in Section II. The bi-level optimization model for analyzing is founded in Section III. The model transformation and linearization procedures are given in Section IV. Case studies are carried out in Section V. The paper is concluded in Section VI.

**II. MODEL OF OAA C OPERATING MODE**

The configuration of the proposed OaaC scheme is firstly demonstrated in Fig. 1. The OaaC participants consist of an operator and a group of users, including residential, small commercial, and industrial users. All of them share the same distribution transformer and are charged by the grid company according to the same rate structure, where both TOU price and demand charge are considered in this study.

**Metrics:**

- $SOC_{\text{min}}$: Minimum allowed SOC of the operator
- $T$: Number of time slots in the optimization horizon
- $T_d$: Number of time slots in a day
- $\eta_c, \eta_d$: Users’ ESS charging/discharging efficiencies
- $H_c, H_d$: Operator’s ESS charging/discharging efficiencies

**Variables:**

- $C$: Metric of violation of the implicit constraint
- $L$: Linearization error

---

**Notation:**

- $D$: Number of time slots in the optimization horizon
- $d$: Number of time slots in a day
- $c$: Users’ ESS charging/discharging efficiencies
- $\eta$: Users’ ESS charging/discharging efficiencies
- $T$: TOU price and demand charge
- $\text{SOC}_\text{min}$: Minimum allowed SOC of the operator
- $\eta_c, \eta_d$: Users’ ESS charging/discharging efficiencies
- $H_c, H_d$: Operator’s ESS charging/discharging efficiencies

---

**Reference:**

1. [Bibliographic entries for references 1 to 9]

---

**Section:**

- Section III: Model formulation
- Section IV: Model transformation and linearization procedures
- Section V: Case studies
- Section VI: Conclusion
This indicates that the operator is equivalent to an ordinary electricity consumer from the grid company’s perspective, namely this scheme showing the characteristic of “operator as a consumer”. Necessary information infrastructure enables bidirectional communication among the operator, the grid, and multiple users.

### A. Interactions Between OaaC Participants

The interactions between the users, the operator, and the grid can be illustrated in Fig. 1 and Fig. 2.

1) **The users:** The users are followers in the Stackelberg game. In the OaaC mode, the participant users send orders of ESS capacities including energy capacity (kWh) orders and power capacity (kW) orders to the operator, as shown in Fig. 1. The amounts of ordered capacities are decided by the users themselves individually. Once ordered, a user can schedule its ordered ESS using its own ESS dispatching method, as if a behind-the-meter ESS was physically deployed. The resulted charging and discharging power orders are sent to the operator. When users send charging power orders, no extra energy is drawn from the grid. Only the State-of-Charge (SOC) of their ordered ESS are modified according to the amount of charging power order. The users pay the energy cost to the operator instead of the grid company for the charging power orders, as depicted in Fig. 2. On the other hand, when a discharging power order is generated by a user, it receives the same amount of energy from the operator, as shown in Fig. 2, and of course its associated SOC is recorded. It has to be noted that the users’ decisions on sending charging/discharging power orders are made by themselves. Meanwhile, the grid company calculates the demand charge for the users based on their shifted load profiles.

2) **The operator:** The operator is the leader in the Stackelberg game. It holds a large-scale shared physical ESS facility and receives ESS orders from the users, as shown in Fig. 1. The detailed behavior of the operator is shown in Fig. 2. The operator sets unit prices of the users’ ESS orders, namely, prices per kWh energy capacity and per kW power capacity, which are also the vital link between the operator and the selfish users, to influence and control the users’ behavior. When discharging power orders are received, the operator has to deliver the users with the required amount of power, provided that their SOCs are not depleted. The power can be either supplied by the installed physical ESS or be purchased by the operator from the grid, depending on its own scheduling strategy. Meanwhile, the operator receives energy cost fees from the users who place charging power orders, and updates their SOCs accordingly.

It should be noted that the users’ ESS orders are essentially some information or data sent to the operator. Neither the operator’s physical ESS is partitioned and assigned to the users, nor the operation of the operator’s physical ESS is directly decided by the users. The key point is that the operator promises to exactly give the users energy for free according to their discharging power orders. The operator’s ESS is used to provide the users’ requested discharged energy. Theoretically, without the operator’s ESS, the OaaC mode can still be deployed; however, the operator has to purchase energy from the grid to fulfill the users’ power needs, which will severely hurt the profitability of the operator.

Moreover, the operator itself is also regarded as a commercial consumer by the grid company, indicating that the electricity price rate including TOU price and demand charge is also enforced. Thus, the operator has the flexibility of not dispatching its ESS exactly according to the charging/discharging power orders from the users. However, how to determine the unit prices for the users’ ESS orders and manage the operator’s shared ESS facility to make this scheme beneficial to both operator and users remains a challenge for further investigation.

### B. Objectives of the Participants

1) **The users:** The objective of the users remains the same as usual, which is to minimize their electricity expense. In this study, it includes energy fees and demand charge paid to the grid company, the costs of their ESS orders, and charging fees paid to the operator. It has to be noted that the users can
conduct their own sizing and scheduling algorithms without releasing their real-time load information.

2) The operator: The objective of the operator can be described as maximizing its net profit, namely the revenue minus the expense. The revenue comes from the users’ payments for their ESS orders and their charging power orders. The expense consists of the investment costs involved by the shared ESS facility and its own electric bill, including energy fees and demand charge.

A detailed model for analyzing the feasibility and economic benefits of the OaaC mode will be given in Section III.

III. PROBLEM FORMULATION

After introducing the working principle of OaaC mode, it remains an interesting question if both the operator’s profit and the users’ savings can be guaranteed, compared with the non-sharing scenario. To address this issue, offline analyses of the proposed OaaC scheme will be carried out to investigate its feasibility and economic performances, i.e. to check whether it is possible that the users’ cost savings and the operator’s profit be achieved simultaneously.

A. Problem Overview

To run the full OaaC scheme, at the start of each billing cycle, i.e. month, the prices for users’ ESS orders are firstly set by operator. Then, OaaC users conduct optimization procedures to decide their ESS orders. After that, across the whole month, at the start of each hour, users decide their charging/discharging orders, and operator optimizes its own ESS operation with those orders as prior knowledge.

Offline analyses are chosen to be given in this paper since full online decision-making algorithms are quite much content and thus deserve another paper. Firstly, an ideal optimization model is formulated without considering any online uncertainties. The global optimum is also obtained. Then, by adding online uncertainties in simulations, the “robustness” of offline optima is numerically analyzed, which will be shown in Section V.

A bi-level optimization model is formulated to characterize the Stackelberg game between selfish operator and users. Bi-level optimization is a common approach for analyzing Stackelberg games [30]. The upper-level problem is the profit-maximization problem of the “leader” operator, where the problems of the “followers” users act as constraints. The lower-level problems are the cost-saving problems of the users, where the operator’s decisions are regarded as known parameters.

B. Lower-Level Model: User Cost Minimization

The lower-level problem is the optimal ESS sizing and scheduling problem for all the users. The objective function for user $i$ is given as its total expense, including electricity bill and costs of ESS orders, over the optimization horizon:

$$\min_{Y_i^t} o^i = c_p, \text{cap} p_{\text{cap}}^i + c_e, \text{cap} e_{\text{cap}}^i$$

$$+ \sum_{t=1}^{T} r(t)(p_L^i(t) + p_C^i(t) - p_D^i(t)) + r_m p_{\text{max}}^i, (1)$$

where $i$ is the index of users, $t$ is the index of time slots, $o^i$ is the objective function of the $i$-th user, $c_p, \text{cap}$ and $c_e, \text{cap}$ are unit prices for ESS power and energy capacity orders, respectively, $r(t)$ is the TOU price, $r_m$ is the demand charge rate, $p_L^i(t)$ is the base load profile, and $Y_i^t = [p_L^i(t), p_C^i(t), p_{\text{cap}}^i, e_{\text{cap}}^i, p_{\text{max}}^i, e^i(t)]$ denotes the decision variables – charging and discharging power orders, power and energy capacity orders, peak power demand, and energy stored in ordered ESS, respectively. The first two terms in (1) are costs of ESS orders. The third term is the energy bill paid to the grid company and the operator. The fourth term is the demand charge.

In the demand charge scheme, a user is charged based on its maximum demand over a billing cycle, usually, a month [1]. The maximum demand is usually chosen as the maximum averaged demand over an averaging interval, e.g. 15min, 30min, or an hour. Hence, the length of each time slot and the length of the optimization horizon $T$ can be chosen as the averaging interval and billing cycle of the demand charge scheme, respectively.

In addition, the constraints are modeled as (2)-(12) below, where $\lambda$ and $\mu$ are dual variables for constraints:

The energy conservation law is:

$$e^i(t) = e^i(t-1) + \eta_p p_{L^i}(t) - \frac{p_{D^i}(t)}{\eta_d}; \quad \mu^i(t), \quad t = 1...T, \quad (2)$$

where $\eta_p, \eta_d$ denote charging/discharging efficiencies, respectively.

The charging and discharging power orders cannot exceed the limits:

$$0 \leq p_{L^i}(t) : \lambda_{L^i}(t), \quad t = 1...T, \quad (3)$$

$$0 \leq p_{D^i}(t) : \lambda_{D^i}(t), \quad t = 1...T, \quad (4)$$

$$p_{L^i}(t) \leq p_{\text{cap}}^i : \lambda_{L^i}(t), \quad t = 1...T, \quad (5)$$

$$p_{D^i}(t) \leq p_{\text{cap}}^i : \lambda_{D^i}(t), \quad t = 1...T. \quad (6)$$

The ordered ESS cannot be charged and discharged in the same time slot:

$$p_{L^i}(t)p_{D^i}(t) = 0, \quad t = 1...T. \quad (7)$$

The peak power demand is calculated as follows:

$$p_{L^i}(t) + p_{C^i}(t) - p_{D^i}(t) \leq p_{\text{max}}^i : \lambda_{D^i}(t), \quad t = 1...T. \quad (8)$$

The user is not allowed to sell electricity back to the grid:

$$0 \leq p_{L^i}(t) - p_{D^i}(t) : \lambda_{D^i}(t), \quad t = 1...T. \quad (9)$$

The stored energy of the ordered ESS cannot exceed the limits:

$$soc_{\text{min}}^i e_{\text{cap}}^i \leq e^i(t) : \lambda_{L^i}(t), \quad t = 0...T, t \neq mT_d, m = 0...D, \quad (10)$$

$$e^i(t) \leq soc_{\text{max}}^i e_{\text{cap}}^i : \lambda_{D^i}(t), \quad t = 0...T, t \neq mT_d, m = 0...D, \quad (11)$$

where $soc_{\text{min}}^i$ and $soc_{\text{max}}^i$ are the lower and upper SOC limits of ordered ESS, respectively, and $D$ is the number of days in the optimization horizon.

At the beginning of each day, the SOC must be the same:

$$e^i(mT_d) = soc_{\text{init}}^i e_{\text{cap}}^i, \quad m = 0...D. \quad (12)$$
where $T_d$ denotes the number of time slots in a day, $soc^i_{ini}$ is the initial SOC of ordered ESS, and $m$ is the index of days.

It can be observed that, equation (7) does not need to be explicitly included if $0 < \eta_c, \eta_d < 1$ and $r(t) > 0$ hold for all $t$. Given $p^c_e(t)$ and $p^d_e(t)$ that $p^c_e(t)p^d_e(t) > 0$, a feasible solution with a lower or same value of objective $o^i$ can be formulated by simply decreasing both $p^c_e(t)$ and $p^d_e(t)$.

Therefore, for each user $i$, the lower-level problem can be modeled as the following Linear Programming (LP) model:

$$
\begin{align*}
\min_{Y^i} & \quad o^i \text{ in (1)} \\
\text{s.t.} & \quad (2) - (6), (8) - (12). \quad (13)
\end{align*}
$$

### C. Upper-Level Model: The Operator’s Profit Maximization

The upper level problem is the operator-side problem that optimizes unit prices for the users’ ESS orders as well as the size and charging/discharging schedule of the shared ESS facility. The operator is committed to providing ESS sharing service to the users at optimal prices, and maximizing its profit (minimizing its net expenditure). The objective function is the total net cost:

$$
\begin{align*}
\min_{Y_o} & \quad O = C_{P, \text{cap}}P_{\text{cap}} + C_{E, \text{cap}}E_{\text{cap}} \\
& \quad + \sum_{t=1}^{T} r(t) \left( P_c(t) - P_d(t) + \sum_{i=1}^{N} p^d_i(t) \right) \\
& \quad + r_m P_{\text{max}} - \sum_{t=1}^{T} \sum_{i=1}^{N} r(t)p^c_i(t) \\
& \quad - C_{P, \text{cap}} \sum_{i=1}^{N} p^c_i - C_{E, \text{cap}} \sum_{i=1}^{N} e^i_{\text{cap}}, \quad (14)
\end{align*}
$$

where

$Y_o = \left[ P_c(t), P_d(t), P_{\text{cap}}, E_{\text{cap}}, P_{\text{max}}, E(t) \right],
\begin{align*}
C_{P, \text{cap}}, C_{E, \text{cap}}, P_{\text{cap}}, E_{\text{cap}}.
\end{align*}
$

denotes the decision variables where $P_c(t), P_d(t)$ are physical charging and discharging power of operator, respectively, $P_{\text{cap}}, E_{\text{cap}}$ are physical power and energy capacities of operator’s ESS, respectively, $P_{\text{max}}$ is peak power demand of operator, and $E(t)$ is stored energy in operator’s ESS. $C_{P, \text{cap}}, C_{E, \text{cap}}$ are unit prices of physical ESS power and energy capacities, respectively. $N$ is the number of users. $O$ is operator’s objective function.

The first two terms in (14) are ESS investment costs. The third term is the energy bill. The fourth term is the demand charge. The fifth term is fees paid by the users for their charging power orders. The rest two terms are the revenue from the users’ ESS orders. In (14), all the decision variables of the lower-level model (13) act as parameters. The lower-level model itself also acts as a constraint in the upper-level model.

The constraints are modeled as (16)-(28) below:

Equations (16)-(23) are similar to the constraints in the lower-level model:

$$
E(t) = E(t-1) + H_c P_c(t) - \frac{P_d(t)}{H_d}, \quad t = 1...T \quad (16)
$$

where $H_c, H_d$ are charging and discharging efficiencies, respectively,

$$
\begin{align*}
0 \leq P_c(t) & \leq P_{\text{cap}}, \quad t = 1...T, \quad (17) \\
0 \leq P_d(t) & \leq P_{\text{cap}}, \quad t = 1...T, \quad (18) \\
P_c(t) & P_d(t) = 0, \quad (19)
\end{align*}
$$

$$
\begin{align*}
P_c(t) - P_d(t) + \sum_{i=1}^{N} p^d_i(t) & \leq P_{\text{max}}, \quad t = 1...T, \quad (20)
\end{align*}
$$

where $SOC_{\text{min}}, SOC_{\text{max}}, SOC_{\text{ini}}$ are lower limit, upper limit, and initial value of SOC of operator’s physical ESS, respectively.

For the operator’s shared ESS facility, there also exist charging/discharging power limits of the battery, which are usually proportional to ESS energy capacity:

$$
\begin{align*}
P_c(t) & \leq k_c E_{\text{cap}}, \quad t = 1...T, \quad (24) \\
P_d(t) & \leq k_d E_{\text{cap}}, \quad t = 1...T, \quad (25)
\end{align*}
$$

where $k_c, k_d$ are charge/discharge rates, respectively.

The peak load at the distribution transformer must not increase after adopting the OoaC mode:

$$
\begin{align*}
P_c(t) - P_d(t) + \sum_{i=1}^{N} p^d_i(t) & \leq \max_{t} \left( \sum_{i=1}^{N} p^d_i(t) \right), \quad t = 1...T. \quad (26)
\end{align*}
$$

Note that it is assumed that the capacity of the distribution transformer is larger than the peak power of aggregated base load profiles. Therefore, the limit of transformer capacity can be neglected with constraint (26) in presence.

Each user’s total expenditure must not increase when participating in the OoaC plan, compared to deploying individual ESS:

$$
o^i \leq f^i, \quad (27)
$$

where $f^i$ is the optimal cost of $i$-th user when using non-sharing individual ESS.

The lower-level problem (13) is also a constraint here:

$$
\begin{align*}
Y^i_t & \in \arg \min_{Y^i_t} \quad o^i, \quad i = 1...N \\
& \quad (2) - (6), (8) - (12). \quad (28)
\end{align*}
$$

Here, the upper-level problem is formulated as:

$$
\begin{align*}
\min & \quad O \text{ in (14)} \\
\text{s.t.} & \quad (16) - (28), \quad (29)
\end{align*}
$$

which is also the bi-level joint optimization problem of ESS pricing, sizing, and scheduling.
IV. TRANSFORMATION INTO MILP

In this section, an equivalent single-level MILP model of the bi-level problem (29) is formulated by two steps: MPEC formulation and linearization. In the MPEC formulation step, the bi-level model is transformed into an equivalent single-level MPEC formulation via various linearization techniques, taking advantage of different model characteristics. This process is concluded in Fig. 3. The derived MILP model can be solved by various kinds of mature algorithms. Therefore, the convergence proof of the solving process is omitted for brevity.

A. MPEC Formulation

By replacing the lower-level problem (28) with its KKT conditions, the bi-level problem (29) can be transformed into an MPEC. The KKT conditions are described below:

1) Primal constraints: (2)-(6), (8)-(12).
2) Dual constraints: Dual constraints are derived by letting the derivatives of the Lagrange function to all the decision variables be zero:

\[ 0 = \frac{\partial L}{\partial p^i_d(t)} = r(t) - \frac{\mu^i(t)}{\eta_d} - \lambda^i_d(t) + \lambda^i_d(t) - \lambda^i_d(t) + \lambda^i_d(t), \quad t = 1...T, \]  
\[ 0 = \frac{\partial L}{\partial p^i_{\text{cap}}(t)} = r_m - \sum_{i=1}^{T} \lambda^i_d(t), \]  
\[ 0 = \frac{\partial L}{\partial e^i_{\text{cap}}(t)} = c_{p, \text{cap}} - \sum_{i=1}^{T} \lambda^i_d(t) - \sum_{i=1}^{T} \lambda^i_d(t), \]  
\[ 0 = \frac{\partial L}{\partial e^i_{\text{cap}}(t)} = c_{p, \text{cap}} - \sum_{i=1}^{T} \lambda^i_d(t) - \sum_{i=1}^{T} \lambda^i_d(t), \]  
\[ 0 = \frac{\partial L}{\partial e^i_{\text{cap}}(t)} = \mu^i(t) - \mu^i(t + 1) - \lambda^i_d(t) + \lambda^i_d(t), \quad t = 0...T, \quad t \neq mT_d, \quad m = 0...D, \]
\[ \lambda^i_d(t) \geq 0, \quad i = 1...N, \quad j = 1...6, \quad t = 1...T, \]
\[ \lambda^i_d(t) \geq 0, \quad i = 1...N, \quad j \in \{7, 8\}, \quad t = 0...T, \quad t \neq mT_d, \quad m = 0...D. \] (37)

3) Complementary slackness: The complementary slackness conditions state that the product of any primal inequality constraint and its dual variable should equal to zero:

\[ 0 = p^i_d(t) \lambda^i_d(t), \quad t = 1...T, \]
\[ 0 = p^i_{\text{cap}}(t) \lambda^i_d(t), \quad t = 1...T, \]
\[ 0 = (p^i_d(t) - p^i_{\text{cap}}(t)) \lambda^i_d(t), \quad t = 1...T, \]
\[ 0 = (p^i_d(t) - p^i_{\text{cap}}(t)) \lambda^i_d(t), \quad t = 1...T, \]
\[ 0 = (p^i_d(t) + p^i_{\text{cap}}(t) - p^i_{\text{max}}(t)) \lambda^i_d(t), \quad t = 1...T, \]
\[ 0 = (p^i_d(t) - p^i_{\text{cap}}(t)) \lambda^i_d(t), \quad t = 1...T, \]
\[ 0 = (p^i_d(t) - p^i_{\text{cap}}(t)) \lambda^i_d(t), \quad t = 0...T, \quad t \neq mT_d, \quad m = 0...D, \]
\[ 0 = (p^i_d(t) - p^i_{\text{cap}}(t)) \lambda^i_d(t), \quad t = 0...T, \quad t \neq mT_d, \quad m = 0...D, \] (45)

Hence, denote the decision variables as

\[ Y = [p^i_d(t), p^i_{\text{cap}}(t), p^i_{\text{max}}(t), e^i_{\text{cap}}(t), e^i_{\text{cap}}(t), p_s(t), p_d(t), \]
\[ p^i_{\text{max}}, p^i_{\text{cap}}, E(t), E_{\text{cap}}, c_{p, \text{cap}}, c_{e, \text{cap}}, \mu^i(t), \lambda^i_d(t)] \] (46)

The MPEC model is formulated as follows:

\[ \min_Y \quad O \quad \text{in } (14) \]
\[ \text{s.t.} \quad (2)-(6), (8)-(12), (16) - (27), (30) - (45), (47) \]

B. Linearization for Tractable Analysis

The MPEC model (47) is still nonlinear and non-convex. Thus, it is difficult to solve. The nonlinearity comes from three sources: constraint (19), complementary slackness conditions, and bilinear terms $c_{p, \text{cap}} p^i_{\text{cap}}, c_{e, \text{cap}} e^i_{\text{cap}}$ in $O$ and $\theta$. In this subsection, an equivalent MILP model is derived by employing multiple linearization methods. Taking advantage of different model characteristics, these methods simultaneously achieve both shorter computing time and enough linearization accuracy.
1) Dealing with (19): (19) can be directly linearized using the big M method [35]. However, this will introduce a large number of binary variables and computing costs. Hence, taking advantage of the complementarity of $P_c(t)$ and $P_d(t)$, a novel model linearization method is proposed and used here to deal with the piecewise-linearity without introducing any integer variables.

Firstly, some variable substitutions are given below:

$$P_b(t) = \begin{cases} H_d P_c(t), & P_c(t) > 0 \\ -P_d(t), & P_d(t) > 0 \\ 0, & P_c(t) = P_d(t) = 0 \end{cases}, \quad t = 1...T, \quad (48)$$

where $P_b$ is the “useful” battery power of the operator, i.e. power that can be finally transmitted to the users, and $E_b$ is the “useful” energy stored by the operator.

Hence, variables $P_c, P_d, E$ can be eliminated if they occurred alone. The corresponding constraints (16), (17)-(19), (21), (22), (23) and (24)-(25) can be rewritten into the following forms, respectively:

$$E_b(t) = E_b(t-1) + P_b(t), \quad t = 1...T, \quad (50)$$

$$-P_{\text{cap}} \leq P_b(t) \leq H_d H_d P_{\text{cap}}, \quad t = 1...T, \quad (51)$$

$$0 \leq \sum_{i=1}^{N} p_d^i(t) + P_b(t), \quad t = 1...T, \quad (52)$$

$$H_d \text{SOC}_{\min} E_{\text{cap}} \leq E_b(t) \leq H_d \text{SOC}_{\max} E_{\text{cap}}, \quad t = 0...T, \quad t \neq mT_d, \quad m = 0...D, \quad (53)$$

$$E_b(mT_d) = H_d \text{SOC}_{\min} E_{\text{cap}}, \quad m = 0...D, \quad (54)$$

$$-k_d E_{\text{cap}} \leq P_b(t) \leq H_c H_d k_c E_{\text{cap}}, \quad t = 1...T. \quad (55)$$

The next work is to eliminate $P_c, P_d, E$ when they occur together. In the studied problem, the term $P_c(t) - P_d(t)$ still lies in the objective function (14) and constraints (20), (26). By involving a slack variable $P_s(t)$, this term can be relaxed into the following formulations:

$$P_c(t) - P_d(t) = P_s(t) + P_b(t), \quad t = 1...T, \quad (56)$$

$$0 \leq P_s(t), \quad t = 1...T, \quad (57)$$

$$\left(\frac{1}{H_d H_d} - 1\right)P_b(t) \leq P_s(t), \quad t = 1...T. \quad (58)$$

The rest work is to prove that this relaxation is exact. Given $P_c(t)$ and $P_b(t)$ that $P_s(t) > 0$ and $P_s(t) > (1/H_c H_d - 1)P_b(t)$, a feasible solution with a lower or same value of objective $O$ can be formulated by simply decreasing $P_s(t)$ until it equals to $(1/H_c H_d - 1)P_b(t).$ This guarantees the exactness of the relaxation.

Therefore, by replacing (16)-(19), (21)-(25) with (50)-(55), and adding constraints (56)-(58), the nonlinear constraint (19) is eliminated without involving any integers.

2) Dealing with the complementary slackness conditions:

Given that the lower-level model (13) is linear, the complementary slackness conditions can be replaced with the strong duality condition [32], stated as below:

$$\sum_{t=1}^{T} (r(t) + \lambda_{o}^{b}(t) - \lambda_{o}^{d}(t)) p_{L}^{i}(t) = o^{i}. \quad (59)$$

where the left side is the dual objective function, denoted by $g^{i}$.

3) Linearizing $c_{p,\text{cap}} p_{\text{cap}}^{i}$ and $c_{e,\text{cap}} e_{\text{cap}}^{i}$: These nonlinear terms lie in the objective functions $O$ and $o^{i}$. It is expected that linear estimations of these nonlinear terms $w_{p}^{i} \approx c_{p,\text{cap}} p_{\text{cap}}^{i}$ and $w_{e}^{i} \approx c_{e,\text{cap}} e_{\text{cap}}^{i}$ are used in these functions, instead of the nonlinear terms themselves. Typical linearization approaches include McCormick relaxation and its piecewise version.

However, for problem (47), the demand for linearization accuracy is so strict that these methods are no longer helpful. The linearization error in a single McCormick envelope may be large according to [36]. When the terms $c_{p,\text{cap}} p_{\text{cap}}^{i}$ and $c_{e,\text{cap}} e_{\text{cap}}^{i}$ are replaced by slack variables $w_{p}^{i}$ and $w_{e}^{i}$, the linearization error may result in deviations of other decision variables’ values according to (1) and (59). This may cause an inaccurate optimal solution, and even infeasible solutions by violating the implicit constraint (7). Hence, to achieve an acceptable linearization accuracy, a large partition number is required in the piecewise McCormick relaxation, leading to a heavy computational burden.

Recently, Normalized Multiparametric Disaggregation (N-MDT) [37] is proposed as a novel piecewise linearization method, which can achieve extremely high accuracy using limited binary variables. By using this discretizing approach, logarithmic growth in the number of binary variables with the partition number is achieved, which grows slowly as compared to the linear growth when using piecewise McCormick linearization [37]. Furthermore, RNMDT [38] is proposed for further reduction of integer variables by using binary expansion instead of decimal system in NMDT. Hence, RNMDT allows us to obtain the same accuracy with far less binary variables introduced.

The basic idea of RNMDT is a combination of nonuniform piecewise linearization and binary expansion. The steps are illustrated in Fig. 4. The generation procedure of RNMDT constraints will be given below; for more detailed steps, please refer to [38]. For brevity purpose, only the linearization of $c_{p,\text{cap}} p_{\text{cap}}^{i}$ is presented for example and $c_{e,\text{cap}} e_{\text{cap}}^{i}$ can be taken care of similarly.

**Step 1. Decompose $c_{p,\text{cap}}$:**

Firstly, it is reasonably assumed that both $c_{p,\text{cap}}$ and $p_{\text{cap}}^{i}$ are bounded. These bounds are denoted by $c_{p,\text{cap},L}, c_{p,\text{cap},H}, p_{\text{cap},L}^{i}, p_{\text{cap},H}^{i}$. Then, $c_{p,\text{cap}}$ can be decomposed into the following form:

$$c_{p,\text{cap}} = c_{p,\text{cap},L} + \lambda_{p}(c_{p,\text{cap},H} - c_{p,\text{cap},L}), \quad (60)$$

where $\lambda_{p} \in [0, 1]$. Furthermore, $\lambda_{p}$ can be represented as follows using binary expansion:

$$\lambda_{p} = \sum_{l \in Z} 2^{l} \cdot z_{p,l}, \quad (61)$$
where \( l \) means the \(-l\)-th (binary) decimal place, and \( z_{p,l} \) are binary variables indicating the number on \(-l\)-th decimal place. For example, \( \lambda_p = (0.101)_2 \) means that \( z_{p,-1} = 1, z_{p,-3} = 1 \), and otherwise \( z_{p,j} = 0 \).

Substituting (61) into (60) gives:

\[
c_{p,\text{cap}} = c_{p,\text{cap},L} + \left( \sum_{l \in \mathbb{Z}_-} 2^l \cdot z_{p,l} \right) \left( c_{p,\text{cap},H} - c_{p,\text{cap},L} \right). \tag{62}
\]

**Step 2.** Cut off the infinite sums:

In practical applications, obviously, in (61) and (62) it is impossible to compute the infinite sums over all negative integers. Hence, a cut-off method by introducing a slack variable \( \Delta \lambda_p \) and the cut-off digit \( Q \leq -1 \) is adopted:

\[
\lambda_p = \sum_{l = Q}^{-1} 2^l \cdot z_{p,l} + \Delta \lambda_p, \tag{63}
\]

\[
0 \leq \Delta \lambda_p \leq 2^Q, \tag{64}
\]

which also brings linearization error, denoted by \( \Delta \lambda_p \). For instance, given \( Q = -2 \), \( z_{p,-1} = 1 \), and \( z_{p,-2} = 0 \), the reconstructed \( \lambda_p \) can be any value in \([0.101]_2, [0.111]_2\], depending on \( \Delta \lambda_p \). A sufficiently small \( \Delta \lambda_p \) can be easily achieved by simply decreasing \( Q \).

Similarly, substituting (63) into (60) gives:

\[
c_{p,\text{cap}} = c_{p,\text{cap},L} + \left( \sum_{l = Q}^{-1} 2^l \cdot z_{p,l} + \Delta \lambda_p \right) \left( c_{p,\text{cap},H} - c_{p,\text{cap},L} \right). \tag{65}
\]

**Step 3.** Give the expression of \( w^i_p \):

Replacing \( c_{p,\text{cap}} \) with the right-hand side of (65) results in the following estimation \( w^i_p \) of \( c_{p,\text{cap}}^i \):

\[
w^i_p \approx c_{p,\text{cap}}^i \begin{cases} 
p^i_{\text{cap}} c_{p,\text{cap},L} \\
+ \left( \sum_{l = Q}^{-1} 2^l \cdot z_{p,l} + \Delta \lambda_p \right) p^i_{\text{cap}} (c_{p,\text{cap},H} - c_{p,\text{cap},L}), \end{cases} \tag{66}
\]

where nonlinear terms \( z_{p,j} p^i_{\text{cap}} \) and \( p^i_{\text{cap}} \Delta \lambda_p \) exist. Since \( w^i_p \) is defined as a linear estimation, these nonlinear terms have to be linearized. Denote \( \tilde{p}^i_{\text{cap},l} \) and \( \Delta v^i_p \) as linear estimations of them, respectively. Then, (66) can be rewritten as:

\[
w^i_p \approx c_{p,\text{cap}}^i \begin{cases} 
p^i_{\text{cap}} c_{p,\text{cap},L} \\
+ \left( \sum_{l = Q}^{-1} 2^l \cdot \tilde{p}^i_{\text{cap},l} + \Delta v^i_p \right) (c_{p,\text{cap},H} - c_{p,\text{cap},L}), \end{cases} \tag{67}
\]

which is linear. The required linearization methods will be given in the following steps.

**Step 4.** Linearize \( z_{p,j} p^i_{\text{cap}} \):

The following constraints are given to ensure that \( \tilde{p}^i_{\text{cap},l} \approx z_{p,l} \tilde{p}^i_{\text{cap}} \), i.e. \( \tilde{p}^i_{\text{cap},l} \) is a linear estimation of \( z_{p,l} \tilde{p}^i_{\text{cap}} \) [38]:

\[
z_{p,l} \tilde{p}^i_{\text{cap},l} \leq \tilde{p}^i_{\text{cap},l} \leq z_{p,l} \tilde{p}^i_{\text{cap},H}, \quad l = Q..-1, \tag{68}
\]

\[
(1 - z_{p,l}) \tilde{p}^i_{\text{cap},L} \leq \tilde{p}^i_{\text{cap},l} \leq (1 - z_{p,l}) \tilde{p}^i_{\text{cap},H}, \quad l = Q..-1, \tag{69}
\]

**Step 5.** Linearize \( p^i_{\text{cap}} \Delta \lambda_p \):

The McCormick relaxation is utilized here to ensure \( \Delta v^i_p \approx p^i_{\text{cap}} \Delta \lambda_p \); i.e. \( \Delta v^i_p \) is a linear estimation of \( p^i_{\text{cap}} \Delta \lambda_p \). The McCormick envelopes are present as follows:

\[
p^i_{\text{cap},L} \Delta \lambda_p \leq \Delta v^i_p \leq p^i_{\text{cap},H} \Delta \lambda_p, \tag{70}
\]

\[
(p^i_{\text{cap}} - p^i_{\text{cap},H}) \cdot 2^Q + p^i_{\text{cap},H} \Delta \lambda_p \leq \Delta v^i_p, \tag{71}
\]

\[
\Delta v^i_p \leq (p^i_{\text{cap}} - p^i_{\text{cap},L}) \cdot 2^Q + p^i_{\text{cap},L} \Delta \lambda_p. \tag{72}
\]

By the aforementioned means, all nonlinear terms are linearized. To sum up, formulas (64)-(65), (67)-(72) are NMDT constraints that need to be added to the single-level optimization model.

**Lemma 1:** The linearization errors \( c_{p,\text{cap}}^i p^i_{\text{cap}} - w^i_p \) and \( c_{e,\text{cap}}^i e^i_{\text{cap}} - w^i_e \) converge to zero as \( Q \) approaches \( -\infty \).

**Proof:** The proof follows a similar procedure as in [37], [38] and is omitted here for brevity.
By applying these techniques, an MILP problem is finally attained in the following and can be solved using mature optimization software, e.g. CPLEX:

\[
\begin{align*}
\min \quad & O = C_{P, \text{cap}} P_{\text{cap}} + C_{E, \text{cap}} E_{\text{cap}} \\
& + \sum_{t=1}^{T} r(t) \left( P_c(t) - P_d(t) + \sum_{i=1}^{N} p_d^i(t) \right) \\
& + r_m P_{\text{max}} - \sum_{i=1}^{N} \sum_{t=1}^{T} r(t) p_i^d(t) - \sum_{i=1}^{N} \left( w_i^f + w_i^j \right),
\end{align*}
\]

s.t. \( (2) - (6), (8) - (12), (27), (30) - (37), \)

\( (50) - (59), (64) - (65), (67) - (72), \)

and RNMDT constraints for \( c_{e, \text{cap}} \).

(73)

where \( Y_M = [Y, E_0, P_b, P_e, w_1^p, z_p, i, \hat{p}_i] \), \( \Delta \lambda_p, \Delta v_p^e, \) \( z_{e, e}, i, \hat{e}_{c, c}, \) \( \lambda_p^e, \Delta v_p^{le}, \) and \( e_{c, c} \) are RNMDT variables involved in the linearization of \( c_{e, \text{cap}} \).

\( \hat{e}_{c, c} \), Note that the objective function in (73) is separable in \( i \). Therefore, distributed solving algorithms such as decomposing methods are also available for solving this derived MILP model.

The problem scale grows at a rate of \( O(NT - NQ) \). Therefore, the time complexity of the solving process can be denoted as \( O(f(NT - NQ)) \) where the function \( f \) is decided by the solving algorithm, e.g. exponential function when branch-and-bound algorithm is selected.

V. SIMULATIONS AND CASE STUDIES

In this section, the feasibility and economic performance of the OaaC mode are examined by case studies using the proposed analyzing method.

A. Real Data as Input

For the simulations, actual load profiles from the Pecan Street database [39] and electricity pricing scheme from Austin Energy [40] are used. The load profiles of consumers in Austin, TX during July 1-31, 2018 are used in this study. The TOU price is $0.0899/kWh during peak hours (4 p.m. - 6 p.m.) and $0.05261/kWh during off-peak hours, while the demand charge rate is set at $7.99/kW. The optimization horizon is chosen as 1 month and the length of a time slot is 1 hour, i.e. \( D = 31, T_d = 24, \) and \( T = T_d D = 744 \).

The base unit prices for ESS energy and power capacities are $300/kWh and $100/kW [15], respectively. The economy of scale factor is set to 60%, meaning that the actual prices \( C_{E, \text{cap}} \) and \( C_{P, \text{cap}} \) are 60% of the base prices, namely, $180/kWh and $60/kW, respectively. These prices are then converted to the monthly costs of $2.0771/kWh and $0.6924/kWh by using the converting function in [14] and setting the lifecycle as 10 years and the discount rate as 6%.

\( f^e \) are acquired by solving the lower-level problem alone by setting \( c_{e, \text{cap}}^e \). Other parameters are shown in Table I.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>soc_{max}</td>
<td>95%</td>
<td>soc_{max}</td>
<td>95%</td>
<td>c_{P, \text{cap}}, L</td>
<td>0.139</td>
</tr>
<tr>
<td>soc_{min}</td>
<td>5%</td>
<td>soc_{min}</td>
<td>5%</td>
<td>c_{P, \text{cap}}, H</td>
<td>2.308</td>
</tr>
<tr>
<td>( k_e )</td>
<td>0.2C</td>
<td>( k_e )</td>
<td>0.5C</td>
<td>c_{E, \text{cap}}, L</td>
<td>0.415</td>
</tr>
<tr>
<td>( \eta_{c, \text{cap}} )</td>
<td>0.05261/kWh</td>
<td>( \eta_{c, \text{cap}} )</td>
<td>0.05261/kWh</td>
<td>( Q )</td>
<td>-12</td>
</tr>
</tbody>
</table>

* \( \max(p_i^d(t)) - \min(p_i^d(t)) \). |

B. Case 1: Study of a Community

In this case study, the MILP model (73) is solved to verify the feasibility and economic performance of the OaaC mode, using data of 10 users. For comparison, two other scenarios are also simulated: a) each user deploys its individual ESS without sharing; b) the users do not deploy any ESS (baseline).

For the month, the result shows that the operator benefits from the OaaC mode. The maximum profit of the operator is calculated as $46.1, with the optimal unit prices for ESS orders being $1.598/kW and $3.095/kW. The revenue from the users’ ESS orders is calculated as $125.0. The revenue from the users’ charging power orders is $50.7. The monthly average expense of deploying the shared ESS facility is $62.1. The electricity bill of the operator is $67.5, where the demand charge is $15.2.

The operator’s profit definitely comes from the users’ ESS orders. Meanwhile, it reduces its own electricity bill, hence “protects” its profit, by optimally dispatching the shared ESS. Although the operator is paid by the users for their charging power orders, it has to purchase almost the same amount of energy from the grid to fulfill the users’ discharging power orders, let alone its demand charge.

Fig. 5 shows the original and shifted load curves of the OaaC operator for the first 2 days. It can be inferred from Section II that the original load curve is exactly the sum of the users’ discharging power orders. Therefore, the operator’s physical ESS dispatching can also be viewed as DR actions. This DR action is quite effective since the users’ discharging power orders are usually concentrated at peak hours, which can be dramatically reshaped using ESS. Without this DR action, the operator’s electricity bill will be $172.5, which is a great loss. Thus, this DR action “protects” the operator’s profit.

Further bill savings for the users are also possible. Fig. 6 shows the cost savings of all users, with individual ESS or participating in the OaaC mode, compared to the baseline. It can be seen that OaaC can save more bills for almost all users compared with the mode of individual ESS. This is because that guaranteeing the users’ cost savings is a major concern of the operator, as depicted by (27).

The optimal ESS energy capacity orders of the users are depicted in Fig. 7. Users 1, 4, 7, 8, 9, 10 would like to place larger ESS orders for better peak shaving. It can be observed that the optimal ESS size of user 10 is almost doubled after participating in the OaaC plan. Their cost savings come from better DR performance. For users 2, 3, 5, 6, their ESS energy
capacities remain almost the same, meaning that their cost savings come from lower prices for their ESS orders.

The solving time is about 4 hours on a 40-core server. Firstly, the considered time horizon in offline analyses here is 1 month, while the solving time grows quickly with the increment of T. In online applications where only day-ahead or intraday decisions are to be made, the computational burden will significantly decrease. Secondly, in offline analyses, a server is used to represent both operator and all users, while in online applications each participant has its own controller and makes its own decisions. Moreover, since the length of a time slot is usually 15 to 60 minutes, communication delay can also be tolerated.

All the following case studies are variations on Case 1.

C. Case 2: Study of A Large Community
A verification study on a larger community with 60 users is then conducted, to explore the profitability of the operator with a large user group.

The resulted operator’s profit is $1044.8. The original and shifted load curves of the operator for the first 2 days are demonstrated in Fig. 8. It can be seen that although more discharging power orders are placed at off-peak time slots compared to Fig. 5, the major demand for discharge is still concentrated at peak hours.

The average of the users’ cost savings is 3.33% compared to the no-ESS scenario. For comparison, individual ESS helps the users save 3.17% costs on average. This indicates that the users’ cost savings are still guaranteed.

D. Case 3: Effect of Battery Ageing
The battery State of Health (SOH) decay caused by ageing is an important factor that affects ESS performance. For a battery, the SOH is the ratio of its actual capacity to its nominal capacity. In this case study, the effects of battery ageing are explored by analyzing the operator’s decisions and profits across the whole lifespan of ESS battery.

The SOH decay is a complex long-term effect [41]–[43]. However, by showing that operator’s profitability is always guaranteed across the whole lifespan of battery, the detailed decay curve can be neglected for conciseness. Three representative scenarios in the battery lifespan are studied here: a) new battery (SOH = 100%, same as Case 1); b) medium decay (SOH = 90%); c) end of lifespan (SOH = 80%).

Under these scenarios, the operator’s profit is $46.1, $36.8, and $25.6, respectively. With the decay of SOH, the operator’s profit undoubtedly decreases but keeps positive in the whole lifespan, which shows great robustness.

E. Case 4: Considering Users’ Different Objectives and Load Shifting
In reality, since the users are allowed to use their own ESS dispatching methods, they may use shared ESS together with load shifting, and the dissatisfaction on load shifting may be a concern of the users. Hence, in this case study, two types of users are considered to explore the effects of load shifting:

a) Users that adopt load shifting and ESS together, with dissatisfaction factor added to their own objectives;

b) Users that do not adopt load shifting.

Data of 10 users with elastic loads from the Pecan Street database [39] is used in this study. It is assumed that users 1-5 belong to type a) with their clothes washers, dishwashers, and water heaters treated as elastic loads, and users 6-10 belong to type b). The model of the users’ load shifting in [9] is
adopted here. Moreover, it is assumed that the operator has no information on the users’ load shifting decisions, to guarantee the users’ autonomy and privacy. Hence, the unit prices for the users’ ESS orders, as well as capacities of the operator’s ESS, are fixed at the optimal values without load shifting.

Fig. 9 delivers type a) users’ ESS capacities with and without load shifting. It can be observed that almost all users’ optimal ordered ESS capacities slightly decrease or keep unchanged after adopting load shifting operations. Even for User 2 whose ordered power capacity increases, its total expenditure on ESS orders still falls. This is because elastic loads can fulfill part of the users’ load reshaping demands.

The operator’s profit is calculated as $144.2, which was $162.8 without the users’ load shifting. This indicates good robustness against users’ load shifting operations.

F. Analyses on Load Prediction Error

Load prediction error has the potential of significantly degrading the performances of all types of DR methods. In this part, different DR methods are used as users’ ESS scheduling methods. At the start of each day, users predict their load curves for that day. Then, users choose to dispatch their ordered ESS by either using day-ahead threshold-based dichotomy method aiming at lowering demand charge only [44] or further applying intraday modifications to the dichotomy results using the information of TOU price. However, for the operator, given the advantage that it knows its load at the start of each hour, it is assumed that it can do a second-stage optimization based on day-ahead results.

The data, optimal unit prices, and optimal ESS orders are kept the same. $P_{\text{cap}}$ is doubled to enhance the operator’s ability of dealing with load fluctuations. This increases the operator’s cost by $7.53. Load predictions are generated from certain normal distributions instead of real prediction methods.

Table II shows the operator’s net profit when all users use the day-ahead demand-charge-oriented dichotomy method only, under different error distributions. For each distribution, 100 Monte Carlo simulations are done, and the average result is shown. The operator’s profit grows with the mean of the users’ load prediction. This is because higher load predictions will lead to users’ higher targets of grid-side power according to the dichotomy algorithm, and consequently fewer and smaller discharging orders. Thus, the operator’s profit is better protected. The profit falls with the growth of the variation of load prediction error.

Table III shows these results when a further intra-day modification according to TOU price is applied. Results show that the operator’s profit is much more robust against load prediction error. This conclusion is trivial since that intraday dispatching methods almost always give better results than day-ahead methods.

These results show that the operator’s profit is quite robust against load prediction error.

G. Sensitivity Analyses on Linearization Accuracy

Sensitivity analyses on linearization accuracy are also carried out to demonstrate the solving performance under different cut-off digit $Q$.

As shown in Section IV-B, the linearization accuracy highly affects the solution and may result in a violation of the implicit constraint (7). However, accurate solutions are always accompanied by high computational burdens. Therefore, the effects of different cut-off digits $Q$ are explored here.

Define a metric for the linearization error:

$$Err_l = \frac{1}{N} \sum_{i=1}^{N} \left( |c_{t,\text{cap}}^i p_{\text{cap}}^i - w_t^i| + |c_{e,\text{cap}}^i e_{\text{cap}}^i - w_e^i| \right),$$

and another metric for assessing the violation of (7):

$$Err_e = \sum_{i=1}^{N} \sum_{t=1}^{T} p_e^i(t) p_d^i(t).$$

The $Err_e$ is zero at any $Q \leq -4$. Fig. 10 shows the results of $Err_l$ and solving time under different $Q$. $Err_l$ decreases exponentially with the increment of $Q$ since the number of equivalent partitions grows exponentially. The computation time grows with the increment of $Q$ exponentially as a result of the linearly-growing number of integer decision variables. It can be seen that $Q = -12$ is a good choice by balancing error and computation time, where the linearization error is small enough that the MILP problem and the original problem can be regarded as equal, and the computation time is acceptable.

These numerical results strongly support the point that the OaaC mode can guarantee both users’ cost savings and
operator’s profit, which delivers good feasibility and economic performance.

VI. CONCLUSION

In this paper, a novel Stackelberg-game-based ESS sharing mode, named the OaaC mode, is proposed for consumers under a complex electricity price system including TOU price and demand charge. This mode achieves users’ full autonomy and privacy protection while guaranteeing both users’ cost savings and operator’s profit maximization. The feasibility and economic performance of the OaaC mode are verified using offline analyzing approaches. Numerical results of OaaC performances are obtained by solving the model using real historical data.

A compelling extension of this study is to develop online robust-stochastic ESS pricing, sizing, and dispatching algorithms for real-time operation of the operator, which will be studied in the future. Moreover, a second-stage optimization targeting at grid-side objectives can also be conducted.

REFERENCES
