Applied Energy 228 (2018) 2510–2525

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Coordinating the operations of smart buildings in smart grids

Yang Liu[a], Nanpeng Yu[b,*], Wei Wang[b], Xiaohong Guan[a], Zhanbo Xu[a], Bing Dong[c], Ting Liu[a]

[a] Systems Engineering Institute, MOE KLINNS Lab, Xi'an Jiaotong University, Xi'an 710049, China
[b] Electrical and Computer Engineering, University of California, Riverside, Riverside, CA 92521, USA
[c] Mechanical Engineering, University of Texas at San Antonio, San Antonio, TX 78249, USA

HIGHLIGHTS

- A novel bi-level building demand aggregation and coordination method is proposed.
- Successive subproblem solving method is introduced to alleviate homogeneous oscillations.
- Three-phase optimal power flow based aggregation at the distribution primary feeder level.
- Building electricity cost is reduced while satisfying all distribution operation constraints.

ARTICLE INFO

Keywords:
Smart building
Load aggregation
Demand response
Proactive demand participation
Building cluster coordination
Distribution network

ABSTRACT

With big thermal storage capacity and controllable loads such as the heating ventilation and air conditioning systems, buildings have great potential in providing demand response services to the smart grid. However, uncoordinated energy management of a large number of buildings in a distribution feeder can push power distribution systems into the emergency states where operating constraints are not completely satisfied. In this paper, we propose a bi-level building load aggregation methodology to coordinate the operations of heterogeneous smart buildings of a distribution feeder. The proposed methodology not only reduces the electricity costs of buildings but also guarantees that all the distribution operating constraints such as the distribution line thermal limit, phase imbalance, and transformer capacity limit are satisfied.

1. Introduction

Increasing integration of intermittent renewable energy resources introduces greater variability and uncertainty into the electricity grid [1]. Thus more ancillary services are required in the electricity market to maintain the reliability of the electricity grid [2], which was provided only by fossil-fueled power plants in the past. Due to the Clean Power Plan that encourages less carbon emissions, more demand response (DR) resources are being procured in the electricity market [3]. With the help of the rapid development of information and control technologies, demand response enables electricity consumers to adjust their electricity usage pattern in response to time-varying electricity price signals, incentive payments and/or direct dispatch instructions. Buildings account for a large amount of the total electricity consumption [4] and Heating, ventilating, and air-conditioning (HVAC) systems consume around a half of buildings' electricity consumption [5]. Hence, if the thermal energy storage inherent in the building is properly managed, buildings can provide an enormous amount of demand response services to the electricity grid.

There is a large body of work which studies energy efficient smart building operations. Lu et al. modeled the major components of HVAC systems and their interactions in building and presented global optimization technologies for economic operation [6]. Guan et al. improved building energy efficiency by coordinating and optimizing the operation of various energy sources and loads in microgrid [7]. Xu et al. studied coordinating multiple storage devices with HVAC systems and determined the optimal operating strategy of building energy systems under time-of-use electricity prices [8]. Maasoumy et al. presented a hierarchical control architecture for balancing comfort and energy consumption in buildings based on a simplified, yet accurate model of the temperature within each room of the building [9]. Ma et al. presented a stochastic model predictive control (MPC) for building HVAC systems considering the load uncertainty of each thermal zone [10]. Radhakrishnan et al. proposed a token-based distributed architecture

* Corresponding author.
E-mail addresses: yliu@sei.xjtu.edu.cn (Y. Liu), nyu@ece.ucr.edu (N. Yu), wwang031@ucr.edu (W. Wang), xhguan@xjtu.edu.cn (X. Guan), zbxu@xjtu.edu.cn (Z. Xu), Bing.Dong@utsa.edu (B. Dong), tingliu@mail.xjtu.edu.cn (T. Liu).

https://doi.org/10.1016/j.apenergy.2018.07.089
Received 8 April 2018; Received in revised form 5 July 2018; Accepted 18 July 2018

Nomenclature

Indices

i	node index in building thermal model, with $i = 1, 2, ..., n-m$ for wall nodes and $i = 1, 2, ..., m$ for room nodes
i', k', n'	index of aggregated nodes under the substation for level 2 aggregation, with $i' = 1, 2, ..., N$, $k' = 1, 2, ..., N$, $n' = 1, 2, ..., N$
j	index of smart building/flexible load under the substation, with $j \in \mathscr{J}$
j'	neighboring node index in building thermal model, with $j' \in \mathscr{N}_{w_i}$ for wall neighboring nodes and $j' \in \mathscr{N}_{r_i}$ for room neighboring nodes
k	index of time interval, with $k = t, t+1, ..., t+W-1$
l	index of the bid points in the demand bid curve, with $l = 1, 2, ..., L$
m', p	phase index, with $m' = 1, 2, 3$, $p = 1, 2, 3$
s	index of secondary feeder system under the substation, with $s \in \mathscr{J}^{sec}$
v	index of discrete levels of the FCU's outlet mass flow rate, with $v = 1, 2, ..., V$

Parameters

A_i	the area of wall i (m^2)
A_{win_i}	the total area of window on walls surrounding room i (m^2)
c_a	the specific heat capacity of air (J/(kg °C))
COP	chiller's coefficient of performance
C_{r_i}	the heat capacity of the indoor air in room i (J/K)
C_{w_i}	the heat capacity of wall i (J/K)
$\boldsymbol{d}_k$	the environment disturbances at time interval k
$F^p_{i',k'}$	real power flow limit between node i' and node k' with phase p (kW)
G^i_{rated}	the rated outlet mass flow rate of the i-th FCU (kg/s)
$GSFP^{p-m'}_{i',k'-n'}$	generation shift factor for real power flow of the branch which connects node i' and k' with phase p when power injection is at node n' with phase m'
g_v	the v-th discrete value in set V (kg/s)
h_i	indicator for room i, 0 if no windows, 1 otherwise
$\mathscr{J}$	the set of all flexible loads under the substation
$\mathscr{J}_s$	the set of flexible loads in a secondary feeder system numbered s
$\mathscr{J}^{sec}$	the set of all aggregated loads after level-1 aggregation
L	total number of distinct bid points in the demand bid curve
m	total number of nodes representing the room air temperature
N	total number of the secondary feeder system under the substation
n	total number of nodes in building thermal model
$\mathscr{N}_{r_i}$	the set of neighboring nodes to node r_i (room i)
$\mathscr{N}_{w_i}$	the set of neighboring nodes to node w_i (wall i)
p_{inc}	the price step for the demand bid curve ($/kWh)
p^i_{rated}	the rated power of the i-th FCU (kW)
$\boldsymbol{pr}_e$	the price forecast vector in energy market ($/kWh)
$\overline{\boldsymbol{pr}_e}$	the upper bound of price forecast ($/kWh)
$\underline{\boldsymbol{pr}_e}$	the lower bound of price forecast ($/kWh)
P^{Tran}_s	the transformer's rated capacity of the secondary feeder system s (kW)
$q^k_{in_i}$	the internal heat generation in room i at time interval k (W)
$q^k_{rad_i}$	the solar radiation density on thermal node i at time interval k (W/m^2)
r_i	indicator for wall i, 0 for internal walls, 1 for peripheral walls
$R_{ij'}$	the resistance between node i and its j'-th neighboring node (K/W)
t	the current time interval
T_{ak}	the ambient temperature at time interval k (°C)
$\boldsymbol{T}_{max}$	the upper bound for indoor temperature (°C)
$\boldsymbol{T}_{min}$	the lower bound for indoor temperature (°C)
$\boldsymbol{T}_s$	the stacked vector of T_{s_i}
T_{s_i}	the temperature of the supply air from the FCU into room i (°C)
u_{max}	the maximum mass flow rate of FCU (kg/s)
u_{min}	the minimum mass flow rate of FCU (kg/s)
V	the total number of discrete levels for the outlet mass flow rate of the FCU
W	the predicting window size (96 points representing 24 h in this paper)
α_i	the absorption coefficient of wall i
β_{win_i}	the transmissivity of glass of window in room i
γ	power imbalance limit between phases (kW)
σ	the penalty factor for SSS method
τ	the length of time interval in each stage (15 min)

Variables

$\boldsymbol{e}_{tot}$	the total energy consumption vector of HVAC system (kWh)
$\boldsymbol{e}^j_{tot}$	total energy consumption vector of HVAC system for building j (kWh)
$G^k_{r_i}$, u^i_k	the mass flow rate of the supply air from the FCU into room i at time interval k (kg/s)
$\boldsymbol{Pd}^{(0)}_j$	the energy consumption of demand bid for individual building j under substation $\mathscr{J}$ (kWh)
$\boldsymbol{Pd}^{(1)}_s$	the energy consumption of demand bid for aggregated load s after aggregating all loads under the secondary feeder system $\mathscr{J}_s$ (kWh)
$\boldsymbol{Pd}^{(2)}$	the final aggregated demand bid at the substation node (kWh)
$PD_{n',m'}$	real power of total demand at node n' with phase m' (kW)
$PG_{n',m'}$	real power of generation at node n' with phase m' (kW)
$\widehat{\boldsymbol{Pd}}^{(1)}_{n',m'}[l]$	demand bid quantity of the l-th segment of the price sensitive demand bid curve at node n' with phase m' (kWh)
$\widehat{\boldsymbol{Pg}}_0[l]$	supply offer quantity of the l-th segment of total supply offer curve at substation node (kWh)
$\widehat{\boldsymbol{Pg}}^{(1)}_{n',m'}[l]$	supply offer quantity of the l-th segment of the supply offer curve at node n with phase m' (kWh)
$P^{m'}_{Loss}$	total real power loss at phase m' (kW)
$P^p_{n'}$	net injection of real power at node n' with phase p (kW)
$\boldsymbol{P}_{tot}$	the total power consumption vector of HVAC system (kW)
$T^k_{j'}$	the temperature of the j'-th neighboring node at time interval k (°C)
$T^k_{r_i}$	the indoor air temperature at time interval k (°C)
$T^k_{w_i}$	the surface temperature of wall i at time interval k (°C)
$\boldsymbol{U}$	all the decision variables for joint-optimization
$\boldsymbol{U}_j$	the outlet mass flow rate of the i-th FCU at time interval k for building j (kg/s)
$\boldsymbol{u}_k$	the stacked input vector representing the air mass flow rate of conditioned air into each thermal zone at time interval k
$\boldsymbol{w}$	the price vector of distinct bid points in the demand bid curve ($/kWh)
$\boldsymbol{w}^d_{n',m'}[l]$	demand bid price of the l-th segment of the price sensitive demand bid curve at node n' with phase m' (k"$"/kWh)
$\boldsymbol{w}^g_0[l]$	supply offer price of the l-th segment of the supply offer curve at substation node ("$"/kWh)

$\boldsymbol{w}^{g}_{n',m'}[l]$	supply offer price of the l-th segment of the supply offer curve at node n' with phase m' ($/kWh)
$\boldsymbol{x}_k$	the stacked state vector representing the temperature of the nodes in the thermal network at time interval k
$\boldsymbol{y}_k$	the temperature vector of each thermal zone at time interval k
y^i_k	the temperature of thermal zone i at time interval k
$z^{i,v}_k$	1 if the outlet mass flow rate of the i-th FCU takes g_v, 0 otherwise
$\hat{\lambda}_t$	the price forecast at current time interval t ($/kWh)

Functions

$C^{d}_{n',m'}$	customer utility function at node n' with phase m'
C^{g}_{0}	cost function of fictious generator at substation node
$C^{g}_{n',m'}$	cost function of distributed generator at node n' with phase m'
$\boldsymbol{f}_j$	functions related to inequality constraints of building j
$\boldsymbol{g}_j$	functions related to equality constraints of building j
P_c	the cooling load power function
P_f	the fan power function

for controlling HVAC systems in commercial buildings, which has low deployment cost and is scalable to buildings with more than 300 zones [11]. To reduce the overall operating cost, Afram and Janabi-Sharifi manipulated the temperature set-points of residential building HVAC systems using an MPC based supervisory controller [12]. In addition, occupancy-based control methods for HVAC systems have been well studied. In particular, Dong and Lam designed and implemented a nonlinear MPC which integrated local weather forecasting with occupant behavior detection, and solved it based on the dynamic programming algorithm [13]. Goyal et al. presented experimental evaluation on two occupancy-based control strategies for HVAC systems in commercial buildings and showed that occupancy-based controllers could yield substantial energy savings over the baseline controllers without sacrificing thermal comfort and indoor air quality [14]. Peng et al. used both unsupervised and supervised learning to learn occupants' behavior, and designed a demand-driven control strategy to make cooling systems automatically adapt to occupants' actual energy demand [15].

The existing building energy simulation and control models can be categorized as physics based (white box) models, data-driven (black box) models, and those in between (gray box models) [16]. The white box models can capture the building dynamics well by using detailed physics-based equations. The white box models such as EnergyPlus [17] and TRNSYS [18] can capture the building thermal dynamics with high accuracy. However, they require detailed information of buildings via extensive energy audit and energy survey. Moreover, the simulations with white box models are extremely time-consuming and not appropriate for real-time applications. The gray box models use simplified physical models to simulate the behavior of building energy systems. For example, Resistance and Capacitance (RC) network model is widely used in online building optimal control and demand response applications, in which different buildings are represented by different RC model parameters [16]. The model parameters are identified based on the operation data using statistics or parameter identification methods, such as nonlinear regression [19], global and local search [20], and genetic algorithm-based parameter identification [21]. However, detailed RC model is still very complicated, which makes the parameter identification and state calculation procedure time-consuming. Hence, model reduction techniques are used to simplify the model which sacrificed some accuracy [22]. The black box models, or the data-driven models, capture the relationship between building energy consumption and operation data based on on-site measurements over a certain period. For example, Vaghefi et al. combined a multiple linear regression model and a seasonal autoregressive moving average model to predict the cooling and electricity demand [23]. Recently, with the rapid development of machine learning (ML) technologies, the ML-based data-driven approaches (black box models) have been well studied. Huang et al. proposed an artificial neural network model to predict the temperature change of multi-zone buildings, and proposed an MPC-based method to maintain the comfortable temperature while reducing energy consumption [24]. Yang et al. presented a reinforcement learning model to control building consists of a PV/T array and geothermal heat pumps [25]. Wei et al. formulated the HVAC control as a Markov decision process and developed a deep reinforcement learning based algorithm to minimize the building energy cost and occupants' discomfort [26]. Behl et al. provided a model-based control with regression trees algorithm, which allows users to perform closed-loop control for DR strategy synthesis for large commercial buildings [27]. Smarra et al. proposed a data-driven MPC using random forests, in which the classical regression tree and random forest algorithms were adapted to determine a closed-form expression for the states prediction function [28]. The data-driven models are model free and require no expert knowledge. After model training, black box models need less computation overhead and are much faster than the gray box models during online optimization. However, black box models often require a large amount of training data and long training period. Moreover, when the operating conditions, weather pattern or building structure change, the trained model is often not usable and needs retraining. Therefore, each of these models has its own advantages and disadvantages. In this paper, we choose the simplified RC model, which is analytically tractable.

It is inefficient and impractical to manage millions of smart buildings directly in the electricity market. Thus load aggregation is one of the key requirements for implementing buildings' DR mechanism. There are already lots of studies on load aggregation. One popular aggregation method is the coordinated aggregation method, which aggregates all the loads into one cluster through linear addition and determines the operation schedule by solving the optimization problem on the cluster level. For example, all the loads under the building cluster are considered together and optimized in a decentralized approach [29]. In the demand response aggregation mechanism [30], the electricity sent to each household is determined by solving a conic quadratic mixed-integer problem at the aggregation node. In [31], the particle swarm optimization is performed to determine the operation strategies for all loads under the building cluster. Regarding the building to grid integration frameworks in both [32,33], all buildings under a transmission network node are regarded as a cluster during optimization. However, this method is unsuitable and inaccurate for the situation where loads are distributed in a large distribution network. Another aggregation method is the bottom-up aggregation, which aggregates loads starting with those connected to low-voltage feeders (residential and small commercial loads fed from distribution transformers), and moving upward toward distribution substations [34]. The bottom-up aggregation method has been widely used in industrial and commercial loads for implementing smart grid functions due to its advantages including easy implementation, fast computation, and wide applicability to load types and variations in power demands [35]. However, this method has limited accuracy and is highly dependent on accurate measurements. Moreover, both methods have not considered the network operating constraints during aggregation, which may lead to issues such as voltage violation, equipment overloads and phase unbalance [36].

To ensure reliable operation of the distribution network, the distribution network constraints should not be ignored. Some distribution network operating constraints are considered in the DR management schemes. In [37], the day-ahead prices for all building loads are

calculated based on social welfare maximization while considering the network operational constraints. The original integer programming optimization problem is relaxed into linear programming problem and solved iteratively in a decentralized approach by the alternating direction method of multipliers (ADMM) based algorithm. In [38], the joint building and grid optimization is implemented in a two-level approach. First, each building is optimized to reduce the electricity cost based on forecasted prices and environment information. Then, based on the optimized load profiles, a distribution grid power flow analysis is carried out. In case of security constraint violation, the maximum allowed load is calculated and sent back to buildings. These two steps are performed iteratively until all the network operating constraints are met. However, in aforementioned work, the active power losses on the distribution lines are not modeled. Furthermore, the optimization procedures are performed on all buildings in the distribution network in each iteration, which increases the model complexity. Furthermore, the proposed approach is time-consuming and not suitable for real-time operations. In [39], most of the distribution network operating constraints are taken into account based on linear programming. The linear approximation of all the constraints and load models improves computational efficiency. However, the approximation could result in performance degradation. In summary, there is a lack of robust algorithm which is capable of coordinating the operations of a large number of smart buildings while considering the distribution network operating constraints. There are two challenges in developing such an algorithm. Firstly, the optimization model for a single building can be nonlinear. Thus the building coordination problem can be very complicated when all the buildings in a distribution feeder are considered. Secondly, the optimal power flow problem in the distribution network is non-convex [40], which makes the load aggregation problem non-convex and hard to solve.

To overcome these difficulties, a novel bi-level aggregation methodology is proposed in this paper, which coordinates the operations of smart buildings in smart grids while considering the operating constraints of the distribution network. The main contributions of this paper are listed below.

- A novel bi-level building load aggregation and coordination methodology is proposed, which not only reduces the building electricity costs, but also satisfies the distribution system operating constraints. The development of the bi-level aggregation is inspired by the physical structure of the distribution network.
- In level-1 aggregation, the joint optimization problem is formulated to coordinate the operations of individual buildings subject to transformer maximum capacity constraint. The SSS method is introduced to decompose the mixed integer linear programming problem (MILP) problem into a series of small coordinated MILP subproblems. This method addresses the homogeneous oscillations problem. Furthermore, the level-1 aggregation can be performed in parallel under each secondary feeder system which makes the approach computationally efficient.
- In level-2 aggregation, the three-phase optimal power flow based aggregation algorithm is developed which not only aggregate the demand bids but also satisfy all the distribution operating constraints. To the best of the authors knowledge, this is the first attempt to develop building aggregation algorithm with a three-phase optimal power flow based approach.

The rest of the paper is organized as follows. Section 2 presents an overview of the proposed smart building operation coordination framework. Section 3 presents the individual building energy scheduling algorithm without coordination. Section 4 presents the proposed bi-level aggregation/disaggregation methodology to coordinate the operations of smart buildings. Section 5 demonstrates the effectiveness of the proposed bi-level aggregation approach with comprehensive simulations, and Section 6 concludes the paper.

2. Overview of smart building operation coordination framework

The overall framework of the proposed smart building operation coordination methodology is illustrated in Fig. 1. The proposed framework is an extension of the proactive demand participation scheme [41]. The overall framework can be divided into three parts: transmission system, distribution system, and individual buildings. There are three types of intelligent decision making entities: the independent system operator (ISO) in the transmission system, the distribution system operators (DSOs) in the distribution network, and the building energy scheduling agents (BESAs) in the smart buildings. The high-level operation procedures of the proposed load aggregation/disaggregation algorithms are described as follows.

- **Load aggregation:** The BESA first collects the information of each individual building and sends it to the DSO. Then, the DSO aggregates the smart buildings and all other flexible loads in the distribution network using the proposed bi-level aggregation method. The output of the bi-level aggregation algorithm is a price-sensitive

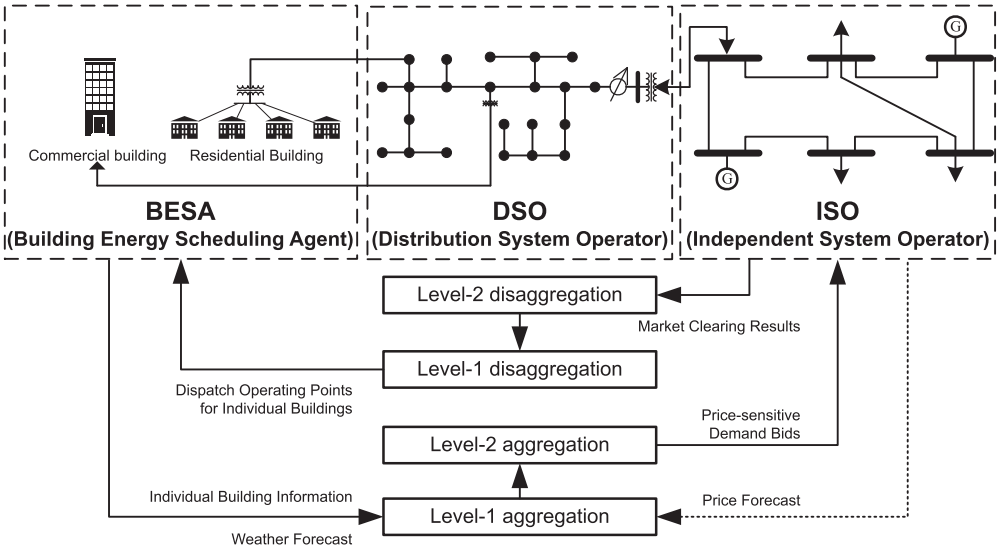


Fig. 1. Coordinated smart building operation framework with bi-level aggregation/disaggregation.

demand bid which represents the overall willingness of all buildings to use electricity under different electricity prices. At this point, the whole distribution feeder/substation is viewed as a virtual power plant [42]. Finally, the DSO sends the aggregated price-sensitive demand bid curve to the ISO for market clearing and resource dispatch.

- **Load disaggregation:** After the wholesale energy market is cleared by the ISO, the market clearing results including the dispatch schedules of the aggregated loads and the locational marginal prices of energy are sent back to the DSO. The DSO then disaggregates the dispatch schedules and sends back the dispatch operating points of individual buildings to each BESA. Finally, the BESA will operate the electrical equipment and follow the dispatch operating points.

To better illustrate the hierarchical bi-level aggregation/disaggregation in the proposed framework, a brief overview of power distribution network topology is given here. A power distribution network can be divided into three levels: the primary feeder, the lateral, and the secondary feeder. At the top level, the primary feeder distributes the electric power from the distribution substation to the lateral feeders. The primary feeder model of the IEEE 13-bus test feeder [43] is shown in Fig. 2a for illustration purpose. In Fig. 2a, node 650 represents the substation node which serves as the point-of-integration to the transmission system. The other nodes can be expanded as the corresponding lateral feeders and secondary feeders as shown in Fig. 2b. The laterals distribute electric power downstream to the secondary feeder systems. Each of the secondary feeders usually consists of the service transformers, the low-voltage secondary lines, and the individual buildings. For example, there are six secondary feeder systems under the sample primary feeder node as shown in Fig. 2b.

Note that since the operating constraints above the secondary feeder system are the same during aggregation and disaggregation, we will consider the aggregation/disaggregation process on the lateral feeder and primary feeder network together without distinction, and classify both of them into the primary feeder level. Thus, we can divide the hierarchical aggregation into two levels and illustrate it as follows.

- **Level-1 aggregation:** The level-1 load aggregation is performed first at the distribution system secondary feeder level (i.e., under the secondary feeder system). A joint optimization problem is formulated for the level-1 aggregation algorithm which takes the transformer capacity constraint into consideration. A sequential subproblem solving method [44] is used to solve the joint optimization problem without homogeneous oscillations.
- **Level-2 aggregation:** The level-2 load aggregation is performed at the distribution system primary feeder level (i.e., above the secondary feeder system) by formulating and iteratively solving a series of three-phase direct-current optimal power flow (DCOPF) problems. The distribution system operating constraints such as the line flow limits, power losses, and the phase imbalance constraints are carefully modeled here.

After the market clears, the disaggregation is performed by DSO based on the dispatched locational marginal price information. Accordingly, the hierarchical disaggregation is divided into two levels, i.e., the level-2 disaggregation at the distribution system primary feeder level and the level-1 disaggregation at the distribution system secondary feeder level. The details of the bi-level aggregation/disaggregation will be illustrated in Section 4.

For any smart buildings/flexible loads being aggregated under the substation, the following hierarchical relationship holds:

$$j \in \mathscr{J}_s \subseteq \mathscr{J}, \quad \forall\, j \in \mathscr{J}, \ \forall\, s \in \mathscr{J}^{sec}$$

where j denotes the subscript for the j-th smart building/flexible load, $\mathscr{J}$ denotes the set of all flexible loads under the substation, and $\mathscr{J}_s$ denotes the set of flexible loads in a secondary feeder system numbered s. After level-1 aggregation, the set of all aggregated loads under each secondary feeder system s is denoted by $\mathscr{J}^{sec}$, and the total number of the secondary feeder system under the substation network is denoted by N. More details about the bi-level aggregation/disaggregation procedure will be illustrated in Section 4.

3. Smart building operation without coordination

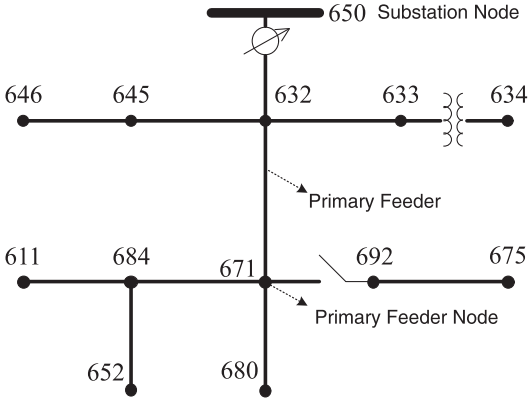
At the individual building level, it has been shown that appropriately managing flexible energy loads can effectively reduce the total energy cost of buildings [45]. In this section, each smart building in the distribution network will be operated without coordination. In the following subsections, we will introduce the building thermal dynamics model, the MPC-based building energy scheduling algorithm, and the demand bid curve generation methodology for an individual building. Finally, all the buildings will be linearly added up based on the generated demand bid curve. For simplification of notations, the subscript for the building index is neglected in this section.

3.1. A model for building thermal dynamics

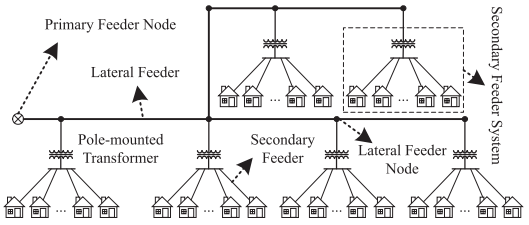
We use the well-established RC networks to model the thermal dynamics of a building as in [9], where each wall and room is modeled as a separate node. After zero-order hold (ZOH) discretization, a discrete-time model representing the building thermal dynamics can be obtained. Suppose there are n nodes in total. m of the nodes represent air temperature in rooms and $n-m$ represent walls. The temperature of the i-th ($i = 1, 2, \ldots, n-m$) wall is governed by the following equation:

$$C_{w_i}(T_{w_i}^{k+1} - T_{w_i}^{k}) = \tau \cdot \left[\sum_{j' \in \mathscr{N}_{w_i}} \frac{T_{j'}^{k} - T_{w_i}^{k}}{R_{ij'}} + r_i \alpha_i A_i q_{rad_i}^{k} \right]$$

where C_{w_i}, α_i and A_i are the heat capacity, absorption coefficient and area of wall i, respectively. $T_{w_i}^{k}$ is the surface temperature of the wall i at time interval k. τ is the length of time interval in each stage. $\mathscr{N}_{w_i}$ is the



(a) Sample primary distribution system.



(b) Sample lateral and secondary feeder system on one primary feeder node.

Fig. 2. One-line diagram for IEEE 13-node test feeder [43] including primary feeders, lateral feeders, and secondary feeder system.

set of neighboring nodes to node w_i (wall i). T_j^k is the temperature of the j'-th neighboring node at time interval k. $R_{ij'}$ is the resistance between wall i and its j'-th neighboring node. r_i is equal to 0 for internal walls and 1 for peripheral walls. $q_{rad_i}^k$ is the solar radiation density on wall i at time interval k.

The air temperature of the i-th ($i = 1, 2, ..., m$) room is governed by the following equation:

$$C_{r_i}(T_{r_i}^{k+1} - T_{r_i}^k) = \tau \cdot \left[\sum_{j' \in \mathcal{N}_{r_i}} \frac{T_{j'} - T_{r_i}}{R_{ij'}} + G_{r_i}^k c_a (T_{s_i} - T_{r_i}) + h_i \beta_{win_i} A_{win_i} q_{rad_i}^k + q_{in_i}^k \right]$$

where C_{r_i} is the heat capacity of the indoor air. $T_{r_i}^k$ is the indoor air temperature at time interval k. $\mathcal{N}_{r_i}$ is the set of neighboring nodes to room i. $G_{r_i}^k$ and T_{s_i} denote the mass flow rate and temperature of the supply air from the fan coil unit (FCU) into room i at time interval k, respectively. c_a is the specific heat capacity of air. A_{win_i} is the total area of window on walls surrounding room i, β_{win_i} is the transmissivity of glass of window in room i. $q_{rad_i}^k$ is the solar radiation density radiated from the window to the room at time interval k, and $q_{in_i}^k$ is the internal heat generation in room i at time interval k. $h_i = 0$ if room i does not have any window, while $h_i = 1$ otherwise.

The above heat transfer differential equations of walls and rooms can be transformed into the following state space equations:

$$\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k + B\boldsymbol{u}_k \circ (\boldsymbol{T}_s - \boldsymbol{y}_k) + E\boldsymbol{d}_k \tag{1}$$

$$\boldsymbol{y}_k = C\boldsymbol{x}_k \tag{2}$$

where the subscript k means the discrete state at time interval k, and $\circ$ is the elementwise product operator for two vectors. $\boldsymbol{x}_k \in \mathbb{R}^n$ is the stacked state vector representing the temperature of the nodes in the thermal network. $\boldsymbol{u}_k = [G_{r_1}^k, G_{r_2}^k, ..., G_{r_m}^k]^T$ is the stacked input vector representing the air mass flow rate of conditioned air into each thermal zone. $\boldsymbol{T}_s = [T_{s_1}, T_{s_2}, ..., T_{s_m}]^T$ is the stacked vector of T_{s_i}. $\boldsymbol{y}_k \in \mathbb{R}^m$ is the temperature of each thermal zone. $\boldsymbol{d}_k$ denotes the environment disturbances.

In the HVAC system, most of the electricity is consumed by the fan, the cooling load and the heating load. Here we assume that heating power is provided by natural gas. Thus the total electricity consumption of HVAC system P_{tot} can be approximated by the sum of fan power P_f and cooling load power P_c. The fan power is approximately proportional to the cubic of its speed. After substituting $G_{r_i}^k$ by $u_k^i = G_{r_i}^k$, the fan power of the i-th FCU can be expressed as follows:

$$P_f(u_k^i) = p_{rated}^i \cdot (u_k^i / G_{rated}^i)^3 \tag{3}$$

where p_{rated}^i and G_{rated}^i are the rated power and rated outlet mass flow rate of the i-th FCU, respectively. The cooling load is estimated by a function of the mass air flow rate and ambient temperature as in [46]:

$$P_c(\boldsymbol{u}_k, \boldsymbol{y}_k) = \frac{c_a}{COP} \sum_{i=1}^{m} u_k^i (T_{ak} - T_{s_i}) \tag{4}$$

where COP is the coefficient of performance for the chiller, and T_{ak} is the ambient temperature at time interval k. So we can obtain the total power consumption of HVAC system at time interval k as follows:

$$\boldsymbol{P}_{tot}[k] = P_c(\boldsymbol{u}_k, \boldsymbol{y}_k) + \sum_{i=1}^{m} P_f(u_k^i) \tag{5}$$

$$\boldsymbol{e}_{tot}[k] = \tau \cdot \boldsymbol{P}_{tot}[k] \tag{6}$$

where $\boldsymbol{P}_{tot}[k]$ and $\boldsymbol{e}_{tot}[k]$ are the total power and energy consumption of HVAC system at time interval k, respectively.

3.2. MPC-based building energy scheduling algorithm

Based on the building thermal dynamics model, we formulate a MPC-based optimization problem in energy market, targeting at minimizing the financial cost while meeting HVAC system's requirements.

$$\min_{u_k^i} \sum_{k=t}^{t+W-1} \boldsymbol{pr}_e[k] \cdot \boldsymbol{e}_{tot}[k] \tag{7}$$

subject to:

$$u_{min} \leqslant u_k^i \leqslant u_{max}, \ \forall\ i, k \tag{8}$$

$$\boldsymbol{T}_{min}[k] \leqslant y_k^i \leqslant \boldsymbol{T}_{max}[k], \ \forall\ i, k \tag{9}$$

Constraints (1), (2), (3), (4), (5), (6)

where $\boldsymbol{pr}_e[k]$ denotes the forecasted locational marginal price at time interval k. u_{min} and u_{max} are the minimum and maximum mass flow rate, respectively. $\boldsymbol{T}_{min}[k]$ and $\boldsymbol{T}_{max}[k]$ are the lower bound and upper bound for indoor temperature at time interval k, respectively. Constraint (8) represents physical bounds for the mass flow rate from FCU. Constraint (9) denotes bounds for room temperature to satisfy building occupants' comfort. Note that the term $\boldsymbol{u}_k \circ \boldsymbol{y}_k$ in Eq. (1) is bilinear, which makes the problem nonconvex and difficult to solve. Since the mass flow rate of FCU is always designed as discrete levels in practice [47], so without loss of generality, we assume that the outlet mass flow rate of the FCU has a total of V discrete levels, and g_v is introduced to indicate the v-th discrete value as in [48]. Thus we have the following equations:

$$u_k^i = \sum_{v=1}^{V} z_k^{i,v} \cdot g_v \tag{10}$$

$$P_f(u_k^i) = p_{rated}^i \sum_{v=1}^{V} z_k^{i,v} \cdot (g_v / G_{rated}^i)^3 \tag{11}$$

$$\sum_{v=1}^{V} z_k^{i,v} \leqslant 1, \quad \forall\ i \tag{12}$$

where $z_k^{i,v}$ is an integer variable, $z_k^{i,v} = 1$ means that the outlet mass flow rate takes g_v at time interval k, otherwise, $z_k^{i,v} = 0$. By substituting constraint (3) with constraint (11), and adding constraints (10) and (12), the original optimization problem can be reformulated as a MILP problem, which can be solved efficiently by commercial solvers such as CPLEX and Gurobi.

The MPC-based algorithm is run periodically. At each time interval t, it determines the optimal mass air flow rate trajectory $[\boldsymbol{u}(t), \boldsymbol{u}(t+1), ..., \boldsymbol{u}(t+W-1)]$ for a predicting window from time t to time $t + W - 1$. Once the optimal trajectory is determined, the MPC-based algorithm will implement the first entry $\boldsymbol{u}(t)$ to control the building HVAC system and operator. Then when the time interval moves forward to $t + 1$, the predicting window will be from time $t + 1$ to time $t + W$.

3.3. Demand bid curve generation for individual buildings

Based on the MPC-based algorithm in previous section, we can construct the price-sensitive demand bid curve for an individual building [41]. For simplification of notation, we use $MPC_j(t, \boldsymbol{pr}_e)$ to present the MPC-based algorithm in Section 3.2 for the j-th building in the distribution network at time interval t based on the price forecast vector $\boldsymbol{pr}_e$. This algorithm will return the j-th building's optimal demand schedules within the predicting window starting at time interval t.

Algorithm 1. Demand bid curve generation for individual building j.

Input: $\boldsymbol{pr}_e$, MPC_j ▷ Other auxiliary inputs: $\overline{\boldsymbol{pr}_e}$, $\underline{\boldsymbol{pr}_e}$, p_{inc}, t
Output: $\boldsymbol{Pd}_j^{(0)}$ ▷ $j \in \mathscr{J}_s$
1: $L \leftarrow (\overline{\boldsymbol{pr}_e}[t] - \underline{\boldsymbol{pr}_e}[t])/p_{inc} + 1$ ▷ Determine number of distinct bid points
2: **for** $l \coloneqq 1$ to L **do** ▷ Traverse all price-demand pairs
3: $\hat{\lambda}_t \leftarrow \underline{\boldsymbol{pr}_e}[t] + p_{inc} * (l-1)$
4: $\boldsymbol{pr}_e[t] \leftarrow \hat{\lambda}_t$ ▷ Update price forecast at time interval t
5: $\boldsymbol{P}_{tot}^{*} \leftarrow \text{MPC}_j(t, \boldsymbol{pr}_e)$
6: $\boldsymbol{w}[l] \leftarrow \hat{\lambda}_t$, $\boldsymbol{Pd}_j^{(0)}[l] \leftarrow \boldsymbol{P}_{tot}^{*}[1]$ ▷ Save energy price-demand pair
7: **end for**

As shown in Algorithm 1, the price forecast vector $\boldsymbol{pr}_e$ and other price and time information are required as input. $\hat{\lambda}_t$ is the price forecast in current time interval t. The number of distinct bid points L is determined at line 1. First, $\hat{\lambda}_t$ is set to the lower bound of price forecast $\underline{\boldsymbol{pr}_e}[t]$. Then the current interval's electricity price in real-time profile $\boldsymbol{pr}_e$ is updated with price forecast $\hat{\lambda}_t$ at line 4. Then after MPC-based algorithm finishes at line 5, the possible energy price and the corresponding demand bid are stored into $\boldsymbol{w}$ and $\boldsymbol{Pd}_j^{(0)}$ at line 6, respectively. In each iteration, $\hat{\lambda}_t$ increases by p_{inc} until $\hat{\lambda}_t$ reaches the upper bound of price forecast $\overline{\boldsymbol{pr}_e}$. Finally, for current time interval t, those isolated energy price-demand pairs $\{(\boldsymbol{w}[l], \boldsymbol{Pd}_j^{(0)}[l]) | l = 1, 2, ..., L\}$ are connected sequentially to form the demand bid curve. A sample price sensitive demand bid curve for an individual customer in a specific time interval is shown in Fig. 3. Particularly, this customer will not use any electricity when the energy price goes above 4.7 ¢/kWh, and want to use at most 0.8kWh electricity when the energy price is between 2.2 ¢/kWh and 4.7 ¢/kWh. When the energy price drops below 2.2 ¢/kWh, this customer will use at most 3.37 kWh electricity.

3.4. Linear additive aggregation

As described in Section 2, after generating the bid curves of individual buildings using Algorithm 1, all buildings will be aggregated together in the proactive demand participation scheme. An aggregated bid curve representing the whole distribution substation will be sent to ISO for market clearing process. Without coordinating the operations of smart buildings, demand aggregation can be performed by directly adding up the bid curves of individual buildings in the network as in [41]. For comparison with the proposed bi-level aggregation approach, we will divide this algorithm into two similar sub algorithms. The aggregation algorithm at the secondary feeder level is shown in Algorithm 2, where the input $\boldsymbol{Pd}_j^{(0)}$ denotes the energy consumption of demand bid for individual building j under substation $\mathscr{J}$, and output $\boldsymbol{Pd}_s^{(1)}$ denotes the energy consumption of demand bid for aggregated load s after aggregating all loads under the secondary feeder system $\mathscr{J}_s$. Afterwards, the aggregation at the primary feeder level is shown in Algorithm 3, where $\boldsymbol{Pd}^{(2)}$ denotes the final aggregated demand bid at the substation node.

Algorithm 2. Linear additive aggregation at secondary feeder level.

Input: $\boldsymbol{Pd}_j^{(0)}$ ▷ $j \in \mathscr{J}_s$
Output: $\boldsymbol{Pd}_s^{(1)}$ ▷ $s \in \mathscr{J}^{sec}$
1: $\boldsymbol{Pd}_s^{(1)} \leftarrow [0]_{1\times L}$ ▷ Initialization
2: **for** each $j \in \mathscr{J}_s$ **do** ▷ Traverse all loads in set $\mathscr{J}_s$
3: $\boldsymbol{Pd}_s^{(1)} \leftarrow \boldsymbol{Pd}_s^{(1)} + \boldsymbol{Pd}_j^{(0)}$ ▷ Linearly add up all bids
4: **end for**

Algorithm 3. Linear additive aggregation at primary feeder level.

Input: $\boldsymbol{Pd}_s^{(1)}$ ▷ $s \in \mathscr{J}^{sec}$
Output: $\boldsymbol{Pd}^{(2)}$ ▷ Aggregated bid at substation node
1: $\boldsymbol{Pd}^{(2)} \leftarrow [0]_{1\times L}$ ▷ Initialization
2: **for** each $s \in \mathscr{J}^{sec}$ **do** ▷ Traverse all loads in set $\mathscr{J}^{sec}$
3: $\boldsymbol{Pd}^{(2)} \leftarrow \boldsymbol{Pd}^{(2)} + \boldsymbol{Pd}_s^{(1)}$ ▷ Linearly add up all bids
4: **end for**

Remark 1 (*Necessity of coordinating operations of smart buildings*). Under normal operation conditions, smart buildings can operate without coordination while satisfying the distribution network constraints. However, it is not suitable to perform simple linear building load aggregation nowadays due to the following reasons: (1) As more buildings proactively participate in the electricity market, the load diversity factor in the distribution network may decrease and the coincident peak demand of loads is very likely to increase; (2) The building energy consumption will increase as new electrical appliances are added based on consumers' new requirements; (3) The degradation of devices (e.g., transformers and lines) may decrease the rated capacities. Thus, the network operating constraints cannot be guaranteed if individual buildings determine their own dispatch operating points without coordination. For example, simulation results in [49] show that when HVAC loads are controlled to respond to certain signals, the load diversity among these HVAC loads is lost, and new load peaks are sharper and the magnitude can be very high. Scenarios in [50] show that with high penetration of electric vehicles, uncoordinated charging could lead to distribution transformer overloading.

4. Smart building operation coordination based on bi-level aggregation/disaggregation

To coordinate the operations of smart buildings, the proposed bi-

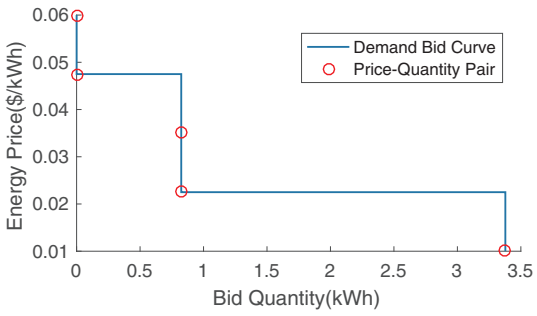


Fig. 3. Sample demand bid curve.

level aggregation and disaggregation approach will be illustrated in details in this section. The distribution network operating constraints are carefully considered in the aggregation and disaggregation process.

4.1. Level-1 aggregation - Demand bid aggregation at the secondary feeder level

In this subsection, a joint optimization based energy scheduling algorithm will be introduced to coordinate the operations of individual buildings subject to transformer maximum capacity constraint. Then, a new demand bid curve generation algorithm is proposed to aggregate the demand bids at the secondary feeder level. Note that the line losses on the secondary feeder system are negligible, because buildings under the secondary feeder system are usually very close to each other.

4.1.1. Joint optimization based building energy scheduling algorithm

To coordinate the operations of smart buildings while satisfying the transformer capacity constraint, a joint optimization algorithm for all the smart buildings under the secondary feeder system is needed. For ease of illustration, the optimization problem in Section 3 for building indexed by j is rewritten as follows.

$$\min_{\boldsymbol{U}_j} \sum_{k=t}^{t+W-1} \boldsymbol{pr}_e[k]\cdot \boldsymbol{e}_{tot}^{j}[k]$$

subject to:

$$\boldsymbol{f}_j(\boldsymbol{U}_j) \leqslant 0,\ \ \boldsymbol{g}_j(\boldsymbol{U}_j) = 0$$

where $\boldsymbol{U}_j = [u_k^{1,j}, u_k^{2,j}, ..., u_k^{m,j}]$, and $u_k^{i,j}$ represents the outlet mass flow rate of the i-th FCU at time interval k for building j, and $\boldsymbol{e}_{tot}^{j}[k]$ represents the total energy consumption at time interval k for building j. $\boldsymbol{f}_j$ and $\boldsymbol{g}_j$ represents all the inequality and equality constraints for building j, corresponding to constraints (1), (2), (4)–(6), (8)–(12). Then by introducing the coupling constraint representing the transformer's capacity limit, we can formulate the following joint optimization problem which is denoted as $\mathrm{MPC}_s^{agg}(t, \boldsymbol{pr}_e)$

$$\min_{\boldsymbol{U}_j} \sum_{k=t}^{t+W-1} \sum_{j\in\mathscr{I}_s} \boldsymbol{pr}_e[k]\cdot \boldsymbol{e}_{tot}^{j}[k] \tag{13}$$

subject to:

$$\boldsymbol{f}_j(\boldsymbol{U}_j) \leqslant 0,\ \ \boldsymbol{g}_j(\boldsymbol{U}_j) = 0,\quad \forall\, j \in \mathscr{I}_s \tag{14}$$

$$\sum_{j\in\mathscr{I}_s} \boldsymbol{e}_{tot}^{j}[k] \leqslant \tau\cdot P_s^{Tran},\quad \forall\, k \tag{15}$$

where the objective function is the summation of individual building's objective function. The constraints of the joint optimization problem includes all individual buildings' operating constraints (14) together with the coupling constraints (15) representing the transformer's capacity limit. P_s^{Tran} is transformer's rated capacity of the secondary feeder system s.

The joint optimization problem is still a MILP problem. The computational complexity of the optimization problem increases exponentially when the number of buildings increases linearly. A Lagrangian relaxation [51] based approach is adopted here to overcome the computational complexity challenge. The coupling constraints (15) is relaxed and added to the objective function by introducing the Lagrangian multipliers. Thus, the joint optimization problem can be divided into subproblems, each for an individual building. However, since the buildings within the same secondary feeder system may have similar optimization parameters, serious solution oscillations may exist when standard Lagrangian relaxation (SLR) based methods are applied. To address the oscillation problem, the successive subproblem solving (SSS) method based on Lagrangian relaxation [44] is applied, which has been used to efficiently solve the unit commitment problem with identical units. The first step of SSS method is to add penalty terms associated with the coupling constraints to the standard Lagrangian as follows:

$$\mathscr{L}(\boldsymbol{U}, \boldsymbol{\lambda}, \sigma) = \sum_{k=t}^{t+W-1} \sum_{j\in\mathscr{I}_s} \boldsymbol{pr}_e[k]\cdot \boldsymbol{e}_{tot}^{j}[k] + \sum_{k=t}^{t+W-1} \lambda_k \left(\sum_{j\in\mathscr{I}_s} \boldsymbol{e}_{tot}^{j}[k] - \tau\cdot P_s^{Tran} \right) + \sigma \sum_{k=t}^{t+W-1} \max\left\{ \sum_{j\in\mathscr{I}_s} \boldsymbol{e}_{tot}^{j}[k] - \tau\cdot P_s^{Tran},\ 0 \right\} \tag{16}$$

where $\boldsymbol{U} = \{\boldsymbol{U}_j \,|\, j \in \mathscr{I}_s\}$ represents all the decision variables, $\boldsymbol{\lambda} = \{\lambda_k \geqslant 0 \,|\, k = t, t+1, ..., t+w-1\}$ denotes the Lagrangian multipliers for relaxing the coupling constraints (15), and $\sigma \geqslant 0$ is the penalty factor. Then, the dual problem can be formulated as follows:

$$\Phi_\sigma^* = \max_{\boldsymbol{\lambda}\geqslant 0} \Phi(\boldsymbol{\lambda}, \sigma)$$

where the dual function is

$$\Phi(\boldsymbol{\lambda}, \sigma) = \min_{\boldsymbol{U}} \mathscr{L}(\boldsymbol{U}, \boldsymbol{\lambda}, \sigma).$$

After introducing the penalty term, the Lagrangian defined in (16) is no longer decomposable. Moreover, the subgradients of $\Phi(\cdot)$ with respect to $\boldsymbol{\lambda}$ are difficult to obtain. However, a surrogate subgradient defined in [52] could be used as a proper direction for updating the multipliers. Define the subproblem for building j as $\min_{\boldsymbol{U}_j} \mathscr{L}_j$ with

$$\mathscr{L}_j(\boldsymbol{U}_j, \boldsymbol{\lambda}, \sigma) = \sum_{k=t}^{t+W-1} \{\boldsymbol{pr}_e[k]\cdot \boldsymbol{e}_{tot}^{j}[k] + \lambda_k \boldsymbol{e}_{tot}^{j}[k]\} + \sigma \sum_{k=t}^{t+W-1} \max\{\boldsymbol{e}_{tot}^{j}[k] + Q_j,\ 0\} \tag{17}$$

where

$$Q_j = \sum_{jj\in\mathscr{I}_s, jj\neq j} \boldsymbol{e}_{tot}^{jj}[k] - \tau\cdot P_s^{Tran}. \tag{18}$$

Since

$$\sum_{j\in\mathscr{I}_s} \boldsymbol{e}_{tot}^{j}[k] - \tau\cdot P_s^{Tran} = \boldsymbol{e}_{tot}^{j}[k] + Q_j \tag{19}$$

By substituting (19) into (16), we have

$$\begin{aligned}\mathscr{L} &= \sum_{k=t}^{t+W-1} \{\boldsymbol{pr}_e[k]\cdot \boldsymbol{e}_{tot}^{j}[k] + \sum_{jj\in\mathscr{I}_s, jj\neq j} \boldsymbol{pr}_e[k]\cdot \boldsymbol{e}_{tot}^{jj}[k]\} \\ &\quad + \sum_{k=t}^{t+W-1} (\lambda_k \boldsymbol{e}_{tot}^{j}[k] + \lambda_k Q_j) + \sigma \sum_{k=t}^{t+W-1} \max\{\boldsymbol{e}_{tot}^{j}[k] + Q_j,\ 0\} \\ &= \mathscr{L}_j + \sum_{k=t}^{t+W-1} \sum_{jj\in\mathscr{I}_s, jj\neq j} \boldsymbol{pr}_e[k]\cdot \boldsymbol{e}_{tot}^{jj}[k] + \sum_{k=t}^{t+W-1} \lambda_k Q_j.\end{aligned} \tag{20}$$

The interaction variable Q_j defined in (18) is related to the dual solutions of the other subproblems, and can be treated as constant if subproblem $\min_{\boldsymbol{U}_j} \mathscr{L}_j$ is solved successively. Thus, the Lagrangian in (20) has two parts: the first term in (20) related to the subproblem for building j, and the other two terms related to the other subproblems.

The SSS method works as follows:

Step 1: **Initialization.** Set the iteration number index $ll = 0$ and initialize $\boldsymbol{\lambda}^0 = 0$. Solve sub-problem $\min_{\boldsymbol{U}_j} \mathscr{L}_j(\boldsymbol{U}_j, \boldsymbol{\lambda}^0, 0)$ iteratively to obtain $\boldsymbol{U}_j^0$ for each building $j \in \mathscr{I}_s$ without the penalty term, which is equivalent to the algorithm in Section 3.2.

Step 2: **Update the Lagrangian multipliers.** Calculate the gradient directions of the Lagrangian multipliers $\boldsymbol{\lambda}$ by

$$g_k^{ll} = \sum_{j\in\mathscr{I}_s} \boldsymbol{e}_{tot}^{j}[k] - \tau\cdot P_s^{Tran} \tag{21}$$

Then choose the step size $step^{ll}$, which satisfies:

$$0 < step^{ll} < (\Phi^* - \mathscr{L}^{ll})/\|\boldsymbol{g}^{ll}\|^2 \tag{22}$$

where Φ^* can be estimated by the method in [44]. Update the Lagrangian multipliers according to:

$$\lambda_k^{ll+1} = \max\{0, \lambda_k^{ll} + step^{ll} \cdot g_k^{ll}\} \quad (23)$$

Step 3: **Update the solutions of the sub-problems.** Find $\boldsymbol{U}^{ll+1}$, which satisfy:

$$\mathscr{L}^{ll+1} = \mathscr{L}(\boldsymbol{U}^{ll+1}, \boldsymbol{\lambda}^{ll+1}, \sigma) < \mathscr{L}(\boldsymbol{U}^{ll}, \boldsymbol{\lambda}^{ll+1}, \sigma) \quad (24)$$

The solutions can be obtained as follows. In terms of building j, substitute the results $\boldsymbol{e}_{tot}^{jj}$ corresponding to $\{\boldsymbol{U}_{jj}^{ll} | jj \in \mathscr{J}_s, jj \neq j\}$ into (18) and solve sub-problem (17) $\min_{U_j} \mathscr{L}_j(\boldsymbol{U}_j, \boldsymbol{\lambda}^{ll}, \sigma)$ to obtain $\boldsymbol{U}^{ll+1}$. If no $\boldsymbol{U}^{ll+1}$ can be found, then let $\boldsymbol{U}^{ll+1} = \boldsymbol{U}^{ll}$. By iteratively solving all sub-problems (17), all the solutions of the sub-problems will be updated while satisfying (24). Note that each subproblem is a small-scale MILP problem and could be solved by MILP solvers efficiently.
Step 4: **Check the criterion.** If $\|\boldsymbol{\lambda}^{ll+1} - \boldsymbol{\lambda}^{ll}\| < \epsilon$ or ll exceeds the maximum allowed iterations, go to step 5; otherwise, go to step 2.
Step 5: **Construct the feasible solution.** The feasible solution is constructed based on the near-optimal solution of the Lagrangian relaxation dual problem obtained by steps 1–4.

The convergence proof of the SSS method can be found in [44]. Since the interaction variable Q_j defined in (18) will be updated after solving each subproblem in Step 3, the dual solutions of buildings with similar optimization parameters could be different. Thus the SSS method could address the homogeneous oscillations associated with traditional Lagrangian relaxation based method.

4.1.2. Demand bid curve generation for level-1 aggregation

To aggregate the building demand with consideration of the transformer's capacity limit, a new demand bid curve generation algorithm for a secondary feeder system is proposed as shown in Algorithm 4. The demand bid curve aggregation algorithm starts by executing the linear additive aggregation Algorithm 2 at line 1. Then the joint optimization model MPC_s^{agg} is formed at line 2. Then we will traverse all the energy price-demand pairs from the lowest price (corresponding to the maximum demand quantity) to the highest price (corresponding to the minimum demand quantity) on the aggregated bid curve $\boldsymbol{Pd}_s^{(1)}$. If the aggregated demand does not exceed the transformer's rated capacity during the iteration process, then linear additive aggregation solution will be selected. Otherwise, the demands bids have to be updated. In such case, we will update the energy price forecast at time interval t corresponding to the price-demand pairs which violate the transformer capacity constraint at line 5. Then we will solve the joint-optimization program MPC_s^{agg} at line 6 using the SSS method. Finally, the energy demand bids $\boldsymbol{Pd}_s^{(1)}[l]$ that violates the constraint will be updated by the coordinated optimal scheduling results $\boldsymbol{P}_{tot}^*[1]$ at line 7. By adopting Algorithm 4, the proposed level-1 demand bid aggregation procedure guarantees that the aggregated demand bids will not violate the transformer's maximum capacity constraint.

Algorithm 4. Level-1 aggregation (Aggregation at secondary feeder level).

Input: $\boldsymbol{pr}_e$, MPC_j, $\boldsymbol{Pd}_j^{(0)}$ ▷ $j \in \mathscr{J}_s$
Output: $\boldsymbol{Pd}_s^{(1)}$ ▷ $s \in \mathscr{J}^{sec}$

1: $\boldsymbol{Pd}_s^{(1)} \leftarrow$ Algorithm 2
2: $\text{MPC}_s^{agg} \leftarrow \text{JOIN}\{\text{MPC}_j | \forall j \in \mathscr{J}_s\}$ ▷ Formulate joint optimization for this secondary feeder system s
3: **for** $l \coloneqq 1$ to L **do** ▷ Traverse all price-demand pairs
4: **if** $\boldsymbol{Pd}_s^{(1)}[l] \geqslant P_s^{Tran}$ **then** ▷ Coupling constraint violated
5: $\boldsymbol{pr}_e[t] \leftarrow \underline{\boldsymbol{pr}_e}[t] + p_{inc} * (l-1)$ ▷ Update price forecast at time interval t
6: $\boldsymbol{P}_{tot}^* \leftarrow \text{MPC}_s^{agg}(t, \boldsymbol{pr}_e)$
7: $Pd_s^1[l] \leftarrow \boldsymbol{P}_{tot}^*[1]$ ▷ Update violated bids
8: **end if**
9: **end for**

4.2. Level-2 aggregation - Demand bid aggregation at the primary feeder level

In this subsection, a three-phase optimal power flow (OPF) based aggregation algorithm is proposed to aggregate the demand bids at the primary feeder level subject to the distribution operating constraints. First, we will introduce the three-phase OPF algorithm, and then we will illustrate how we can aggregate the demand bids up to the substation node. After aggregation, the entire distribution feeder or substation can be treated as a virtual power plant with its own aggregated demand bid curve.

4.2.1. Three-phase optimal power flow algorithm

The key operating constraints which need to be considered in the aggregation process at primary feeder level include the phase imbalance constraints and the line flow limit constraints. Besides, the distribution line losses should be carefully modeled. A three-phase DCOPF model [53] can be leveraged to coordinate the operations of various distributed energy resources while satisfying the operating constraints of the power distribution network. The details of the DCOPF model is provided here.

Suppose there are $N+1$ node in the distribution system, the substation node (for example, node 650 in Fig. 2a) is denoted by node 0, and the other aggregated nodes or renewable distributed resources belonging to $\mathscr{J}^{sec} = \{1, 2, ..., N\}$ are denoted by node 1 to node N. Node 0 is selected as the swing bus of the distribution system, and an equivalent system supply offer is created at node 0. The aggregated demand bids on node 1 to node N will be decomposed into three phases for problem formulation.[1] Then the three-phase DCOPF problem can be formulated as follows:

$$\max \sum_{n'=1}^{N} \sum_{m'=1}^{3} \{C_{n',m'}^d(\widehat{\boldsymbol{Pd}}_{n',m'}^{(1)}) - C_{n',m'}^g(\widehat{\boldsymbol{Pg}}_{n',m'}^{(1)})\} - C_0^g(\widehat{\boldsymbol{Pg}}_0) \quad (25)$$

subject to:

$$\sum_{n'=0}^{N} PG_{n',m'} = \sum_{n=0}^{N} PD_{n',m'} + P_{Loss}^{m'}, \quad m' = 1, 2, 3 \quad (26)$$

$$\left| \sum_{n'=1}^{N} \sum_{m'=1}^{3} GSFP_{i',k'-n'}^{p-m'}(PG_{n',m'} - PD_{n',m'}) \right| \leqslant F_{i',k'}^p, \quad \forall\, i',\ k' \text{ and } i' \neq k',\ p = 1, 2, 3 \quad (27)$$

$$\left| \sum_{n'=1}^{N} P_{n'}^{m'} - \sum_{n'=1}^{N} P_{n'}^{p} \right| \leqslant \gamma,\ m', \quad p = 1, 2, 3 \text{ and } m' \neq p \quad (28)$$

$$C_{n',m'}^d(\widehat{\boldsymbol{Pd}}_{n',m'}^{(1)}) = \sum_{l=1}^{L} \boldsymbol{w}_{n',m'}^d[l] \cdot \widehat{\boldsymbol{Pd}}_{n',m'}^{(1)}[l], \quad \forall\, n' \in \mathscr{J}^{agg} \text{ and } m' = 1, 2, 3 \quad (29)$$

$$C_{n',m'}^g(\widehat{\boldsymbol{Pg}}_{n',m'}^{(1)}) = \sum_{l=1}^{L} \boldsymbol{w}_{n',m'}^g[l] \cdot \widehat{\boldsymbol{Pg}}_{n',m'}^{(1)}[l], \quad \forall\, n' \in \mathscr{J}^{agg} \text{ and } m' = 1, 2, 3 \quad (30)$$

[1] In the previous aggregation procedures, the aggregated bids provision can be calculated on three phases. However, the phase term is omitted for simpler notation. For a single-phase load, the loads on other two phases are regarded as fixed loads with zero demand.

$$C_0^g(\widehat{\boldsymbol{Pg}}_0) = \sum_{l=1}^{L} \boldsymbol{w}_0^g[l]\cdot\widehat{\boldsymbol{Pg}}_0[l]. \tag{31}$$

where $\widehat{\boldsymbol{Pd}}_{n',m'}^{(1)}[l]$ and $\widehat{\boldsymbol{Pg}}_{n',m'}^{(1)}[l]$ represent the demand bid quantity and the supply offer quantity of the l-th segment of the price sensitive demand bid curves at node n' with phase m', respectively. $\widehat{\boldsymbol{Pg}}_0[l]$ denotes the supply offer quantity of the l-th segment of total supply offer curve at substation node. $PG_{n',m'}$ and $PG_{n',m'}$ are the real power of generation and total demand at node n' with phase m', respectively. $P_{Loss}^{m'}$ is the total real power loss at phase m', respectively. $GSFP_{i',k'\_n'}^{p\_m'}$ is the generation shift factor for real power flow of the branch which connects node i' and k' with phase p when power injection is at node n' with phase m'. $P_{n'}^{p}$ is the net injection of real power at node n' with phase p. $F_{i',k'}^{p}$ is the real power flow limit between node i' and node k' with phase p. γ is the power imbalance limit between phases. $\boldsymbol{w}_{n',m'}^{d}[l]$ is the demand bid price of the l-th segment of the price sensitive demand bid curve at node n' with phase m'. $\boldsymbol{w}_{n',m'}^{g}[l]$ is the supply offer price of the l-th segment of the supply offer curve at node n' with phase m'. $\boldsymbol{w}_0^g[l]$ is l-th supply offer price of the l-th segment of the supply offer curve at substation node.

The objective function (25) maximizes the total surplus of customers and producers in a distribution system. The first term of function (25) denotes customers' utility function, the second term denotes the sum of generation cost for each node except for Node 0, while the last term denotes the generation cost of Node 0. The real power balance constraints are represented by Eq. (26). Eq. (27) is the power flow limit constraints which guarantees that the power flow will not exceed the thermal capacity on each distribution line. Phase imbalance constraints are represented in Eq. (28), which are effective in mitigating phase imbalance problems. Customer utility function and generator cost function are calculated in Eqs. (29)–(31).

The three-phase DCOPF problem can be solved by the iterative three-phase DCOPF algorithm in our previous work [53], which is capable of finding a good approximation to the three-phase alternative-current optimal power flow (ACOPF) problem in a computationally efficient manner. Under the assumption of unitary voltage and small angle deviations, the power flow equation could be linearized around the flat solution, and the system parameters including $GSFP_{i',k'\_n'}^{p\_m'}$ for the linearized three-phase power flow equation can be obtained.[2] Then the three-phase DCOPF problem (Eqs. (25)–(31)) can be solved as a linear optimization problem. In our three-phase DCOPF algorithm, the linear optimization problem is solved iteratively until the solution converges. The fictitious nodal demand (FND) model in [54] is adopted in the algorithm, which can distribute system losses among distribution lines to eliminate significant mismatch at the reference bus. As shown in [54], the FND-based DCOPF yields a closer approximation to the results of ACOPF. The three-phase DCOPF algorithm is summarized as follows:

Step 1: Initially set linearized system parameters, power injections and power flows. Set FNDs, power losses to zeros.
Step 2: Solve the linear optimization problem, update the power injections and power flows.
Step 3: Update the parameters of the linearized system, FNDs and power losses based on the new solution.
Step 4: Solve the linear optimization problem again.
Step 5: Check the dispatch of loads and generation resources. If the difference between the current iteration and previous iteration's result is larger than the pre-defined tolerance, go the Step 3. Otherwise, the final three-phase OPF solution is obtained.

With the supply offer price at the substation node and the demand bid curve for the other nodes as inputs, we can easily compute the optimal dispatch operating points for each node based on the proposed three-phase DCOPF algorithm.

[2] Detailed derivations can be found in [53].

4.2.2. Demand bid curve generation for level-2 aggregation

Suppose that the supply offer price at the substation node is fixed at some value (i.e., the supply offer bid curve is a straight line) in the three-phase OPF problem, the dispatch demand quantity at each node indicates how much energy each node want to consume at current offered electricity price. Particularly, the dispatched demand quantity at the substation node indicates how much energy all the buildings under this substation want to consume at a certain electricity price. As we increase (decrease) the supply offer price at the substation node, the corresponding demand quantity at substation node will decrease (increase). These pairs of supply offer price and demand quantity explicitly quantify the flexibility of all loads under this feeder/substation. Based on this idea, we can perform the aggregation procedure at the primary feeder level by Algorithm 5.

Algorithm 5. Level-2 aggregation (Aggregation at primary feeder level).

Input: $\boldsymbol{Pd}_{s,m'}^{(1)}$ ▷ $s \in \mathscr{S}^{sec}$ and $m' = 1, 2, 3$
Output: $\boldsymbol{Pd}^{(2)}$ ▷ Aggregated bid at substation node

1: $L \leftarrow (\overline{\boldsymbol{pr}_e}[m'] - \underline{\boldsymbol{pr}_e}[m'])/p_{inc} + 1$
2: **for** $l := 1$ to L **do**
3: $\quad\lambda_0 \leftarrow \underline{\boldsymbol{pr}_e}[m'] + p_{inc} * (l-1)$
4: $\quad\boldsymbol{w}_0 \leftarrow [\lambda_0]_{1\times L}$ ▷ Set supply offer price
5: $\quad(PG_{0,1}^*, PG_{0,2}^*, PG_{0,3}^*) \leftarrow$ DCOPF $(\boldsymbol{w}_0, \boldsymbol{Pd}_{s,m'}^{(1)})$
6: $\quad\boldsymbol{w}[step] \leftarrow \lambda_0$ ▷ Save energy price
7: $\quad\boldsymbol{Pd}^{(2)}[step] \leftarrow \sum_{m'=1}^{3} PG_{0,m'}^*$ ▷ Save total demand
8: **end for**

As shown in Algorithm 5, for each possible energy bid price λ_0 at line 3, the supply offer price at substation node is set to be constant as that price at line 4. After solving the DCOPF problem at line 5, we can get the demand quantities on each phase at substation node, which also equals to the total optimal dispatched load on each phase, respectively. Finally, at line 6 and line 7, all the price-total demand pairs in each iteration are stored and connected sequentially to form the aggregated demand bid curve on each phase at the substation node. The generated bid curve will be submitted to the wholesale market by the DSO.

4.3. Demand disaggregation

After the wholesale market clears, the dispatch operating points for the aggregated loads need to be disaggregated into the dispatch instructions at the individual building level. Since we aggregate the demand bids at two levels, we will also disaggregate the dispatch operating signals at two levels by using the locational marginal price information.

Level-2 disaggregation - Disaggregation at the primary feeder level: First, with the cleared market price at the substation node, the three-phase DCOPF problem in Section 4.2 is solved again. The marginal price for each primary feeder node can be calculated after solving this DCOPF problem [53].

Level-1 disaggregation - Disaggregation at the secondary feeder level: If the dispatched load for the secondary feeder system do not exceed the transformer's maximum capacity, the price signal is sent to each individual building directly, and the MPC-based algorithm is run separately to determine the optimal schedule for the current time interval. Otherwise, based on the marginal price for each aggregated node presenting a secondary feeder system, the joint optimization model MPC_s^{agg} in Section 4.1 is solved again to determine the dispatched

schedules for each individual buildings. Finally, each individual building will control its flexible loads according to the dispatch schedules.

5. Simulation and analysis

In this section, we investigate the impact of smart building operations on the distribution grid, and demonstrate the effectiveness of the proposed bi-level aggregation methods.

5.1. Simulation setup

Numerical studies are conducted on the IEEE 13-node test feeder [43] as shown in Fig. 2. HVAC control systems are assumed to be the major flexible loads of a typical building model in the simulation. The simulations are implemented in MATLAB on a PC with 3.30-GHz Intel (R) Xeon(R) E3-1226 v3 CPU and 8 GB of RAM. The MILP subproblem is modeled by YALMIP [55] and solved by Gurobi [56]. The major simulation parameters are chosen as follows.

- The ambient temperature and solar radiation for a whole day in the simulation are shown in Fig. 4. The forecasted energy prices are based on PJM's historical price data and are shown in Fig. 5.
- The reference building model is adapted from [46], whose parameters have been validated through EnergyPlus [17] simulation. The building is modeled as a single zone with four peripheral walls, one roof and one floor. The zone size is 10 m× 10 m× 3 m, and the thermal parameters are shown in Table 1.
- There are already lots of occupancy data set available online [57,58]. For simplicity, we use four typical different occupancy patterns to represent the customers' occupancy behaviors in our simulation. As shown in Fig. 6, within each horizontal bar representing 24 h, occupied hours are filled with color, and unoccupied hours are left blank.
- Regarding the level-1 aggregation, we assume that there are ten buildings under each of the secondary feeder system as shown in Fig. 2b. The thermal parameters of each building are randomized around the reference building model, and the occupancy profile of each customer is generated by randomly picking one of the occupancy patterns.
- In terms of the level-2 aggregation, we perform the simulation based on aggregated bid curves on the primary nodes. More details will be shown in Section 5.3.
- For each individual building, the comfortable indoor temperature (defined by the lower temperature bound $\boldsymbol{T}_{min}$ and upper temperature bound $\boldsymbol{T}_{max}$ in (9)) is determined by its predicted occupancy presence. The comfortable indoor temperature should fluctuate between 21 °C and 25 °C when the building is occupied. There are no requirements for indoor temperature when the building is unoccupied.
- Regarding the MPC-based algorithm, the time interval t is assumed to be 15 min, and the predicting window W is set to be 24 h.

5.2. Case study for level-1 aggregation

In this subsection, we will first compare the proposed level-1 building load aggregation algorithm with two benchmarking algorithms, and show that the proposed level-1 building aggregation and coordination algorithm not only reduces building electricity costs but ensures reliable operation of power distribution network. Then we will analyze the SSS method adopted here and demonstrate that it alleviates the homogeneous oscillation problem caused by the standard Lagrangian relaxation method. At last, a typical aggregated demand bid curve will be generated as the output of the level-1 aggregation.

5.2.1. Evaluation of the proposed level-1 aggregation algorithm

The performance of the proposed level-1 building load aggregation and coordination algorithm will be compared with two benchmarking algorithms through three different aggregation scenarios under a secondary feeder system. The transformer's rated capacity for this secondary feeder system is 35 kW. The setup of the three aggregation scenarios are summarized in Table 2. The implementation details of the three scenarios are described here.

Scenario I: In the first benchmarking algorithm, individual buildings do not participate in any demand response program. In addition, buildings do not explicitly coordinate with each other when controlling flexible loads. The bang-bang controller [59] does not optimally control the HVAC because the control input is fixed at the maximum level when the HVAC is turned on [32]. Thus, to make a fair comparison, a multi-state control model similar to [60] is adopted here. In the multi-state control model, different control input levels are triggered at different temperature bands. The temperature bands and the input levels are carefully selected and tuned to make sure the temperature will not violate the temperature bound constraint (9). All building loads under this algorithm will be aggregated by the linear additive aggregation Algorithm 2.

Scenario II: In the second benchmarking algorithm, individual buildings participate in the proactive demand response program without considering the network operating constraints. The MPC-based algorithm in Section 3.2 is utilized to control the HVAC system. All building loads under the secondary feeder are aggregated by the linear additive aggregation Algorithm 2.

Scenario III: In the proposed level-1 building aggregation and coordination Algorithm 4, individual buildings participate in the proactive demand response program while considering the network operating constraints. As described in Section 4.1, the joint optimization model MPC_s^{agg} is formulated and solved by the SSS method to control all HVAC systems under the secondary feeder.

The simulation results of three scenarios are shown in Table 2, and the active power of the aggregated loads of the distribution secondary under the three scenarios are shown in Fig. 7. Compared with scenario I, scenario II achieved a lower building electricity cost (12.48% reduction) through price-based MPC algorithm. However, it also introduced a higher peak load. This is mainly caused by the uncoordinated building load operations. With similar optimization parameters and real-time electricity price forecasts, the load diversity factor under scenario II becomes much lower than that of scenario I. Hence, the peak load of scenario II increased to 38.31 kW which exceeds the transformer's rated capacity constraint. Similar phenomenon is also shown in [49]. In contrast, the proposed algorithm in scenario III coordinated the operations of the smart buildings and reduced the peak load of the secondary feeder below the transformer's rated capacity. Although the electricity bill in scenario III is slightly higher than scenario II, our proposed algorithm still achieved a 9.73% electricity cost reduction when compared with scenario I.

5.2.2. Performance of the SSS method

To make a comparison between the SSS method and the SLR method, the joint optimization problem MPC_s^{agg} is solved by the two

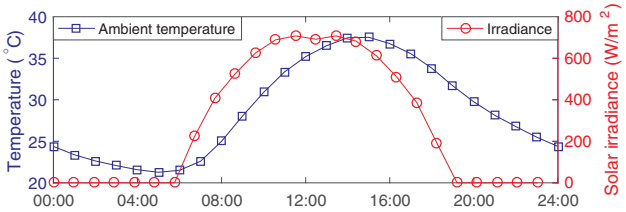


Fig. 4. Ambient temperature and solar irradiance for a whole day in the simulation.

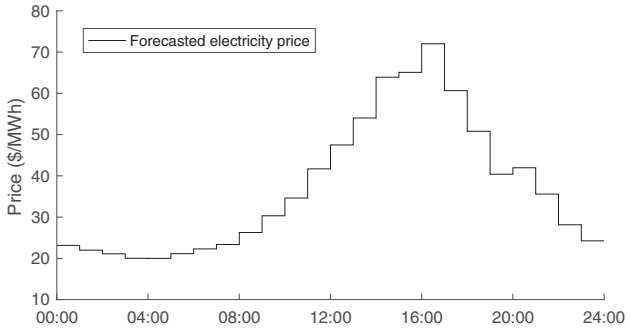


Fig. 5. Forecasted electricity price in the PJM market.

Table 1
Thermal parameters of the reference building model.

Parameter	Value	Definition
C_{w_1}	2.39×10^6 J/K	Thermal capacitance of four peripheral walls.
C_{w_2}	7.89×10^6 J/K	Thermal capacitance of roof/floor.
C_{air}	3.69×10^5 J/K	Thermal capacitance of the room air.
$R_{out,1}$	1.19×10^{-2} m·K/W	Thermal resistance between the peripheral wall and the outside air.
$R_{out,2}$	3.61×10^{-3} m·K/W	Thermal resistance between roof/floor and the outside air.
$R_{in,1}$	1.36×10^{-2} m·K/W	Thermal resistance between the peripheral wall and the room air.
$R_{in,2}$	4.11×10^{-3} m·K/W	Thermal resistance between roof/floor and the room air.

methods respectively. The concept of violation degree [44] are used here to evaluate the convergence rate and the solution feasibility of both methods. The violation degree is defined as $\sum_{k=t}^{t+W-1} \max\{\sum_{j\in\mathcal{J}_s} \boldsymbol{e}_{tot}^{j}[k] - P_s^{Tran}, 0\}$, which represents the total amount of power violation in a whole day. The violation degree measures "how far" a dual solution is away from a feasible one.

As shown in Fig. 8, the feasible solutions and infeasible solutions are marked with circles and asterisks respectively in each of the optimization iterations. From the results, we can see that after the fifth iteration, the SLR method starts oscillating between two solutions (one feasible solution and one infeasible solution) and has a difficult time converging. This is because customers with similar building parameters and occupancy patterns have similar electricity usage behavior. On the other hand, with the help from the additional penalty term, the SSS method can find a feasible solution after one iteration, and converges quickly after the third iteration. Since the coupling constraints are simple and can be met with little efforts, the SSS method can always converge after a few iterations in our simulation. Therefore, the SSS method successfully mitigates the homogeneous oscillation problem for the joint optimization problem MPC_s^{agg} in the level-1 building load aggregation algorithm.

Fig. 8 shows that the SLR method may face the oscillation problem in the level-1 load aggregation algorithm. Hence, the SSS method is adopted in the proposed framework.

Table 2
Main features and results of three different aggregation scenarios.

Scenario #	Features			Results		
	Control method	DR	Coordination	Aggregation method	Peak load (kW)	Total cost ($)
Scenario I	Multi-state	No	No	Algorithm 2	31.98	16.03
Scenario II	MPC	Yes	No	Algorithm 2	38.31	14.03
Scenario III	MPC_{agg}	Yes	Yes	Algorithm 4	34.30	14.47

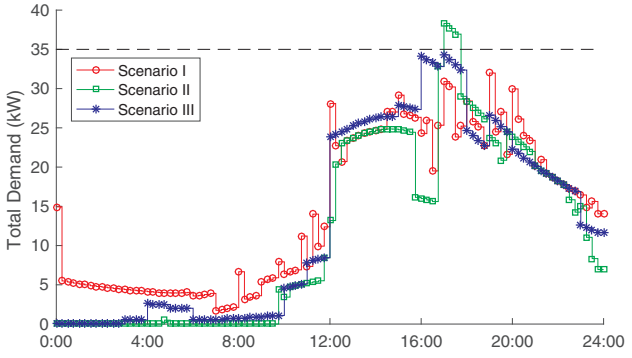


Fig. 7. Total demand of all loads under a secondary feeder under the three scenarios.

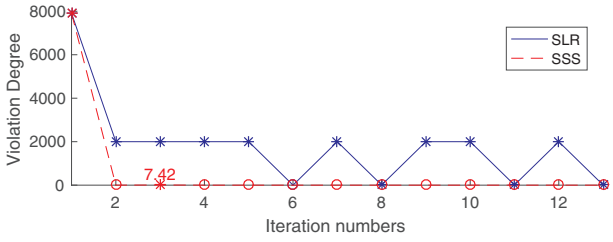


Fig. 8. Violation degree versus dual iteration.

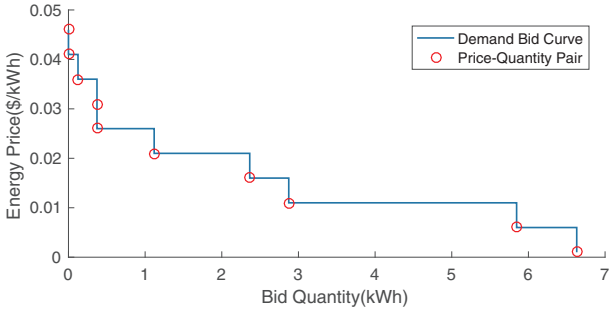


Fig. 9. Demand bid curve for the secondary feeder system at 11:00 a.m. after level-1 aggregation.

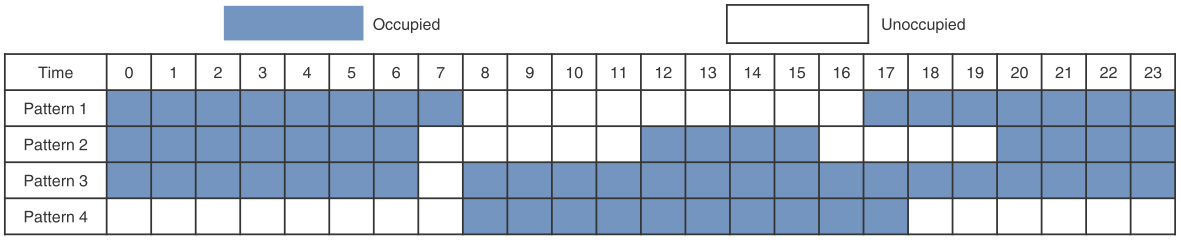


Fig. 6. Buildings' occupancy patterns in the simulation.

5.2.3. Demand bid curve of the level-1 aggregation

After applying the proposed level-1 building load aggregation Algorithm 4, the aggregated demand bid curves for this secondary feeder are generated. For example, the bid curve at 11:00 a.m. is shown in Fig. 9.

5.3. Case study for level-2 aggregation

In this subsection, we will compare the proposed level-2 building load aggregation algorithm with the linear additive algorithm at the primary feeder level. The simulations are set up as follows. Assume that there are four flexible aggregated loads at node 633, 634, 652 and 611. The fixed loads are set up in the same way as in the IEEE 13-bus test feeder benchmark document [43]. It is assumed that the flexible loads of the three phases are not completely balanced at node 633, 634, 652, and 611. The demand bid curve shown in Fig. 9 is used as the bid curve for each secondary feeder to construct the bid curves for the flexible loads. The final demand bid curves for flexible loads on the four nodes are shown in Fig. 10. Note that node 652 and 611 are single-phase nodes.

Two scenarios are simulated. In **scenario A**, the demand bids on each node will be aggregated by the linear additive load aggregation Algorithm 3. In **scenario B**, the demand bids on each node will be aggregated based on the level-2 load aggregation Algorithm 5. Different distribution network constraints are analyzed under both scenarios to demonstrate the effectiveness of our proposed level-2 building load aggregation algorithm.

(1) Analysis of the line flow limit: In the simulation, the thermal limit for line ⟨632–633⟩ in Fig. 2a is set to be 400 kVA. After load aggregation process, the daily maximum apparent power flows on three phases of this line under different bid prices are shown in Fig. 11. As can be seen from the figure, the linear additive load aggregation Algorithm 3 results in thermal limit violation when the energy bid price is lower than \$0.02/kWh. The proposed level-2 load aggregation Algorithm 5 in scenario B, on the other hand, satisfies the thermal limit constraints all the time.

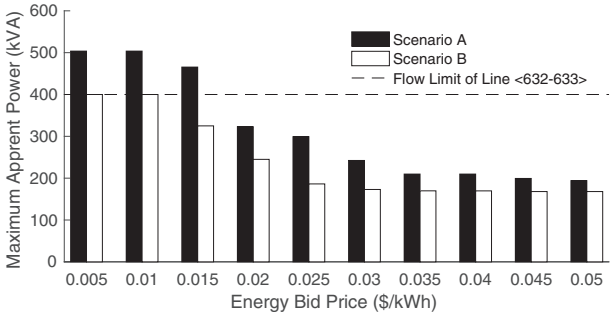


Fig. 11. The maximum apparent power flows on three phases of line ⟨632–633⟩ under different bid prices of the two scenarios.

(2) Analysis of the three-phase imbalance: The maximum allowed phase imbalance power is set to be 60 kW in the simulation. The aggregated demand bid curves of the phase *a*, *b*, and *c*, as well as the maximum phase imbalance at the substation node are presented in Fig. 12. It can be easily seen that the bid curves for three phases in scenario B (Fig. 12b) are much closer to each other than that of the scenario A (Fig. 12a). As shown in Fig. 12c, the linear additive load aggregation Algorithm 3 violates the maximum phase imbalance constraints. In contrast, by utilizing the proposed level-2 aggregation Algorithm 5 in scenario B, the maximum phase imbalance does not exceed the maximum allowed phase imbalance power.

(3) Analysis of the power losses: The linear additive load aggregation Algorithm 3 used in scenario A ignores the power losses in the

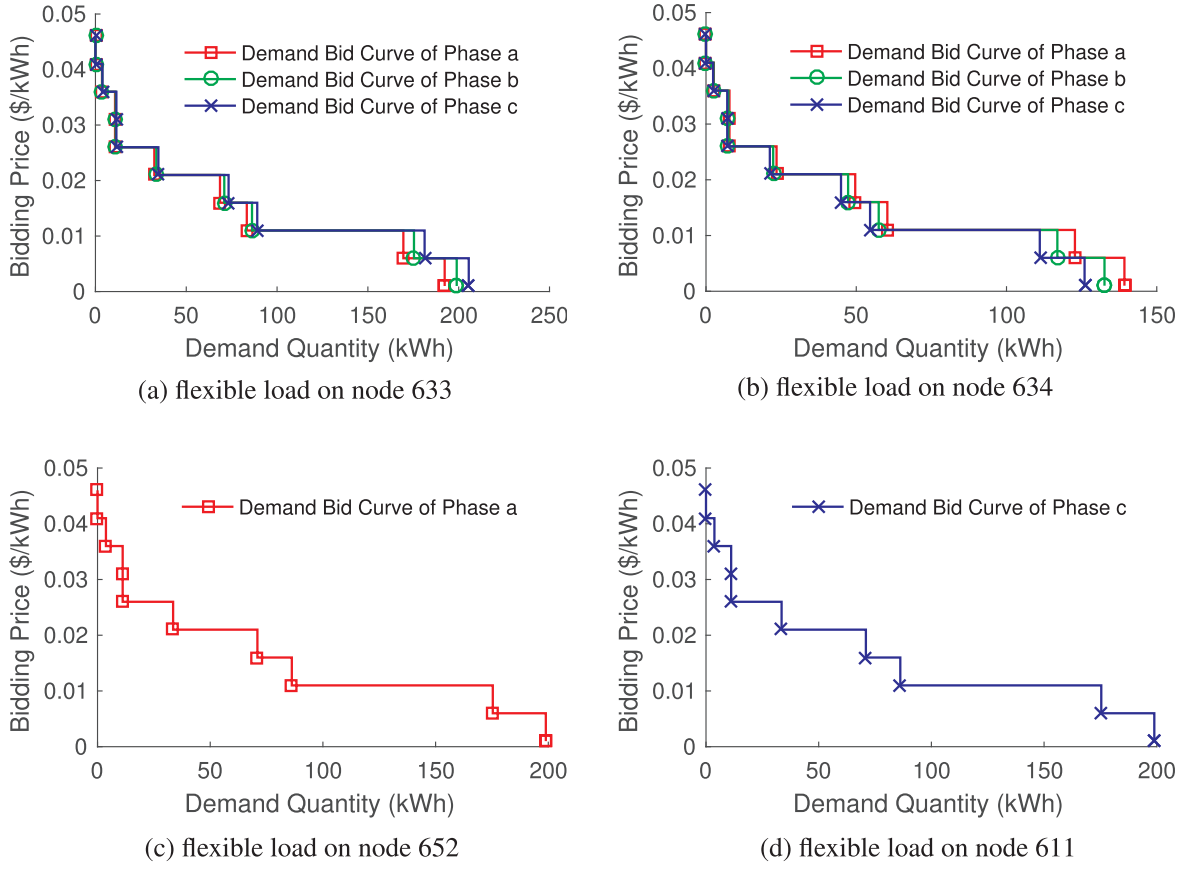
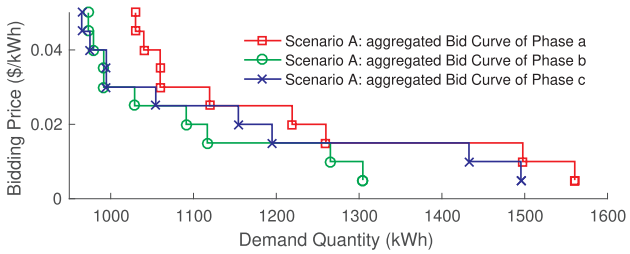
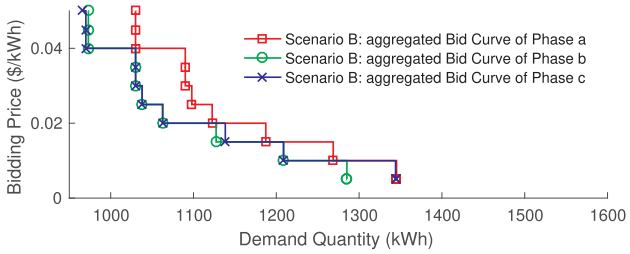


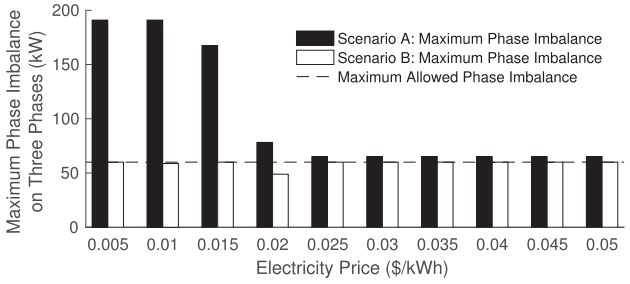
Fig. 10. Bid curves for the flexible loads on node 633, 634, 652 and 611 in IEEE 13-node test feeder.



(a) Aggregated bid curve of phase *a*, *b*, and *c* under scenario A.



(b) Aggregated bid curve of phase *a*, *b*, and *c* under scenario B.



(c) Maximum phase imbalance on three phases under different bid prices.

Fig. 12. Aggregated bid curves of phase, *a*, *b*, *c*, and the maximum phase imbalance at the substation node.

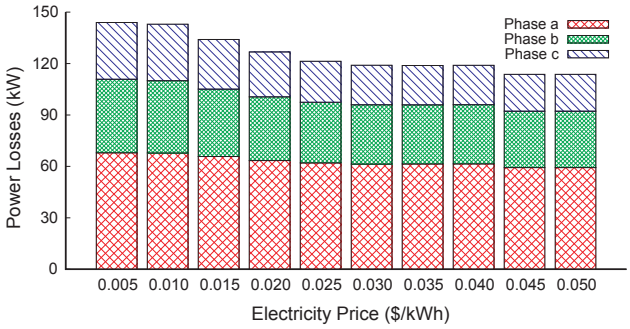


Fig. 13. Power losses under different bid prices in scenario B.

distribution network. In contrast, the proposed level-2 load aggregation Algorithm 5 leverages the iterative three-phase DCOPF algorithm shown in Section 4.2 to capture the power losses in scenario 2. This feature makes the proposed load aggregation much more accurate. As shown in Fig. 13, the total power losses under different bid prices range from 113.7 kW to 144.0 kW, which account for 3–5% of the total demand. Hence, the power losses cannot be ignored and should definitely be considered in the load aggregation process.

5.4. Scalability of the aggregation algorithm

In this subsection, the scalability of the bi-level aggregation algorithm will be verified by test cases of different sizes.

To validate the scalability of the level-1 aggregation algorithm, simulations are conducted by assuming different number of buildings are connected to the distribution transformer in the secondary feeder system. The transformer's rated capacity is chosen in such a way that the peak load without coordination will exceed the rated capacity. As discussed in Section 4.1.1, the subproblems of SSS method in Algorithm 4 are solved sequentially. However, the joint optimization problem under different energy price forecasts (i.e., Step 6 in Algorithm 4) can be solved in parallel. The simulation results are shown in Table 3, where the third column represents the maximum parallel computation time of SSS method under ten different bid prices.

It can be seen from Table 3 that the computation time of the SSS method increases approximately in a linear fashion when the number of buildings under a distribution transformer increases. This is because the subproblems of SSS method are solved sequentially. However, in the real world, the number of buildings connected to a distribution transformer is very limited. For example, a commonly used 25 kilovolt-ampere (kVA) neighborhood transformer serves on average five to seven homes [61]. Another example is that a typical 12 kV distribution feeder [62] serves between 1000 and 2000 customers with over 400 transformers. Thus, the computation time for level-1 aggregation is less than 21 s in most real-world cases.

To validate the scalability of the level-2 aggregation algorithm, simulations are conducted on five IEEE distribution feeder test cases. In each test case, flexible loads are added on 50% of the nodes, and the total computation time for Algorithm 5 is recorded. Note that to create the aggregated demand bid, the DCOPF algorithm runs ten times for ten different bidding prices. As shown in Table 4, the computation time of Algorithm 5 is very short. For the 123-bus system, the total computation time for Algorithm 5 is less than 3 s.

The above simulation results show that the bi-level aggregation algorithm takes less than one minute to finish in most real-world cases. The performance is quite reasonable as the real-time electricity market is cleared every five minutes. In summary, the simulation results showed that the proposed methodology is robust, scalable, and can be implemented in real-time market operations.

It should be noted that the unitary voltage assumption is used in the three-phase DCOPF model. However, there exist some cases where some long, rural feeders which may face severe voltage problem. To take the severe voltage problem into consideration, the ACOPF algorithm proposed in [63], which has already included the voltage constraint, can be adapted to replace the DCOPF algorithm here. Based on the results of [63], the computation time of the ACOPF algorithm on the 123-bus test system is around 27 s. Thus, if we adopt the ACOPF algorithm in parallel under different bid prices in Algorithm 5, the final total computation time can meet operation requirement in the real-time energy market. Meanwhile, the proposed bi-level aggregation framework still works after adopting this change.

6. Conclusion

This paper proposes a novel bi-level building demand aggregation methodology to coordinate the operations of smart buildings in smart grids. The proposed method improves upon the existing work by taking the key distribution system operating constraints including the line thermal limit, phase imbalance, and transformer capacity limit into consideration during the aggregation process. At the distribution

Table 3
Scalability of level-1 aggregation.

Number of buildings	P_{max} (kW)	Maximum computation time (s)
6	21	16.49
10	35	21.07
20	70	43.15
30	105	74.53

Table 4
Scalability of level-2 aggregation.

Test system	Number of flexible loads	Total computation time (ms)
4-bus	2	440.10
13-bus	6	747.50
34-bus	17	597.03
37-bus	19	1129.74
123-bus	60	2650.59

secondary feeder level, a joint optimization problem is formulated to perform the level-1 aggregation. The successive subproblem solving method is introduced to alleviate the homogeneous oscillations problem. At the distribution primary feeder level, a three-phase direct-current optimal power flow based method is developed to perform the level-2 aggregation. The simulation results demonstrate that the proposed smart building coordination and aggregation method not only reduces building electricity costs but also satisfies all distribution system operating constraints.

In the future, we plan to extend the proposed smart buildings aggregation framework in three directions. First, the other types of flexible loads such as stationary energy storage systems and electric vehicles will be incorporated into the modeling framework. Second, we will explore ways to develop a three-phase alternative-current optimal power flow based smart buildings aggregation algorithm at the primary feeder level to better represent the nonlinearity of the distribution networks. Third, we will investigate the tradeoff between the accuracy of the building thermal dynamics model and the complexity of the optimization formulation of the secondary level building aggregation problem.

Acknowledgement

This work was supported by National Key Research and Development Program of China under grant (2016YFB0901905), National Natural Science Foundation of China under grants (61472318, 61632015, 61772408, 6180022135, U1766215, U1736205), National Science Foundation (NSF) under awards (#1637258, #1637249), Department of Energy under award (#DE-OE0000840), Fok Ying Tong Education Foundation (151067), and the Fundamental Research Funds for the Central Universities.

References

[1] Huber M, Dimkova D, Hamacher T. Integration of wind and solar power in Europe: assessment of flexibility requirements. Energy 2014;69:236–46. https://doi.org/10.1016/j.energy.2014.02.109.
[2] D. Wang, D. Kalathil, K. Poolla, X. Guan, Coordination of wind power and flexible load through demand response options, in: IEEE Conference on Decision and Control, 2015, pp. 7226–7231. https://doi.org/10.1109/CDC.2015.7403359.
[3] Siano P. Demand response and smart grids—a survey. Renew. Sustain. Energy Rev. 2014;30:461–78. https://doi.org/10.1016/j.rser.2013.10.02.
[4] U.S. Energy Information Administration, Annual Energy Outlook 2017, January 2017 < https://www.eia.gov/outlooks/aeo/ > .
[5] Hao H, Lin Y, Kowli AS, Barooah P, Meyn S. Ancillary service to the grid through control of fans in commercial building HVAC systems. IEEE Trans. Smart Grid 2014;5(4):2066–74. https://doi.org/10.1109/TSG.2014.2322604.
[6] Lu L, Cai W, Chai YS, Xie L. Global optimization for overall HVAC systems – part I problem formulation and analysis. Energy Convers. Manage. 2005;46(7):999–1014. https://doi.org/10.1016/j.enconman.2004.06.012.
[7] Guan X, Xu Z, Jia Q. Energy-efficient buildings facilitated by microgrid. IEEE Trans. Smart Grid 2010;1(3):243–52. https://doi.org/10.1109/TSG.2010.2083705.
[8] Xu Z, Guan X, Jia QS, Wu J, Wang D, Chen S. Performance analysis and comparison on energy storage devices for smart building energy management. IEEE Trans. Smart Grid 2012;3(4):2136–47. https://doi.org/10.1109/TSG.2012.2218836.
[9] M. Maasoumy, A. Pinto, A. Sangiovanni-Vincentelli, Model-based hierarchical optimal control design for HVAC systems, in: ASME 2011 Dynamic Systems and Control Conference and Bath/ASME Symposium on Fluid Power and Motion Control, 2011, pp. 271–278. https://doi.org/10.1115/DSCC2011-6078.
[10] Ma Y, MatuÅko J, Borrelli F. Stochastic model predictive control for building HVAC systems: complexity and conservatism. IEEE Trans Control Syst Technol 2015;23(1):101–16. https://doi.org/10.1109/TCST.2014.2313736.
[11] Radhakrishnan N, Su Y, Su R, Poolla K. Token based scheduling for energy management in building HVAC systems. Appl Energy 2016;173:67–79. https://doi.org/10.1016/j.apenergy.2016.04.023.
[12] A. Afram, F. Janabi-Sharifi, Supervisory model predictive controller (MPC) for residential HVAC systems: implementation and experimentation on archetype sustainable house in Toronto. Energy Build 154. https://doi.org/10.1016/j.enbuild.2017.08.060.
[13] Dong B, Lam KP. A real-time model predictive control for building heating and cooling systems based on the occupancy behavior pattern detection and local weather forecasting. Build Simul 2014;7(1):89–106. https://doi.org/10.1007/s12273-013-0142-7.
[14] Goyal S, Barooah P, Middelkoop T. Experimental study of occupancy-based control of HVAC zones. Appl Energy 2015;140:75–84. https://doi.org/10.1016/j.apenergy.2014.11.064.
[15] Peng Y, Rysanek A, Nagy Z, Schlter A. Using machine learning techniques for occupancy-prediction-based cooling control in office buildings. Appl Energy 2018;211:1343–58. https://doi.org/10.1016/j.apenergy.2017.12.002.
[16] Li X, Wen J. Review of building energy modeling for control and operation. Renew Sustain Energy Rev 2014;37:517–37. https://doi.org/10.1016/j.rser.2014.05.056.
[17] Crawley DB, Pedersen CO, Lawrie LK, Winkelmann FC. EnergyPlus: energy simulation program. ASHRAE J 2000;42(4):49.
[18] TRNSYS, TRNSYS: Transient System Simulation Tool < http://www.trnsys.com/ > .
[19] Braun JE. Reducing energy costs and peak electrical demand through optimal control of building thermal storage. ASHRAE Trans 1990;96(2):876–88.
[20] ho Lee K, Braun JE. Model-based demand-limiting control of building thermal mass. Build Environ 2008;43(10):1633–46. https://doi.org/10.1016/j.buildenv.2007.10.009.
[21] Wang S, Xu X. Simplified building model for transient thermal performance estimation using GA-based parameter identification. Int J Therm Sci 2006;45(4):419–32. https://doi.org/10.1016/j.ijthermalsci.2005.06.009.
[22] Goyal S, Barooah P. A method for model-reduction of non-linear thermal dynamics of multi-zone buildings. Energy Build 2012;47:332–40. https://doi.org/10.1016/j.enbuild.2011.12.005.
[23] Vaghefi A, Jafari M, Bisse E, Lu Y, Brouwer J. Modeling and forecasting of cooling and electricity load demand. Appl Energy 2014;136:186–96. https://doi.org/10.1016/j.apenergy.2014.09.004.
[24] Huang H, Chen L, Hu E. A neural network-based multi-zone modelling approach for predictive control system design in commercial buildings. Energy Build 2015;97:86–97. https://doi.org/10.1016/j.enbuild.2015.03.045.
[25] Yang L, Nagy Z, Goffin P, Schlueter A. Reinforcement learning for optimal control of low exergy buildings. Appl Energy 2015;156:577–86. https://doi.org/10.1016/j.apenergy.2015.07.050.
[26] Wei T, Wang Y, Zhu Q. Deep reinforcement learning for building HVAC control. In: 2017 54th ACM/EDAC/IEEE design automation conference (DAC); 2017. pp. 1–6. https://doi.org/10.1145/3061639.3062224.
[27] Behl M, Smarra F, Mangharam R. DR-Advisor: a data-driven demand response recommender system. Appl Energy 2016;170:30–46. https://doi.org/10.1016/j.apenergy.2016.02.090.
[28] Smarra F, Jain A, de Rubeis T, Ambrosini D, DInnocenzo A, Mangharam R. Data-driven model predictive control using random forests for building energy optimization and climate control. Appl Energy. https://doi.org/10.1016/j.apenergy.2018.02.126.
[29] Vrettos E, Oldewurtel F, Vasirani M, Andersson G. Centralized and decentralized balance group optimization in electricity markets with demand response. In: PowerTech (POWERTECH), 2013 IEEE Grenoble. IEEE; 2013. p. 1–6. https://doi.org/10.1109/PTC.2013.6652519.
[30] Mhanna S, Verbi G, Chapman AC. Towards a realistic implementation of mechanism design in demand response aggregation. In: Power systems computation conference; 2014. p. 1–7. https://doi.org/10.1109/PSCC.2014.7038379.
[31] Li X, Wen J, Malkawi A. An operation optimization and decision framework for a building cluster with distributed energy systems. Appl Energy 2016;178:98–109. https://doi.org/10.1016/j.apenergy.2016.06.030.
[32] Taha AF, Gatsis N, Dong B, Pipri A, Li Z. Buildings-to-grid integration framework. IEEE Trans Smart Grid 2017(99):1. https://doi.org/10.1109/TSG.2017.2761861.
[33] Dong B, Li Z, Taha A, Gatsis N. Occupancy-based buildings-to-grid integration framework for smart and connected communities. Appl Energy 2018;219:123–37. https://doi.org/10.1016/j.apenergy.2018.03.007.
[34] Gkatzikis L, Koutsopoulos I, Salonidis T. The role of aggregators in smart grid demand response markets. IEEE J Select Areas Commun 2013;31(7):1247–57. https://doi.org/10.1109/JSAC.2013.130708.
[35] Saleh SA, Pijnenburg P, Castillo-Guerra E. Load aggregation from generation-follows-load to load-follows-generation: Residential loads. IEEE Trans Indust Appl 2017;53(2):833–42. https://doi.org/10.1109/TIA.2016.2626261.
[36] Clement-Nyns K, Haesen E, Driesen J. The impact of charging plug-in hybrid electric vehicles on a residential distribution grid. IEEE Trans Power Syst 2010;25(1):371–80. https://doi.org/10.1109/TPWRS.2009.2036481.
[37] Moradzadeh B, Tomsovic K. Two-stage residential energy management considering network operational constraints. IEEE Trans Smart Grid 2013;4(4):2339–46. https://doi.org/10.1109/TSG.2013.2265313.
[38] Razmara M, Bharati G, Hanover D, Shahbakhti M, Paudyal S, Robinett R. Building-to-grid predictive power flow control for demand response and demand flexibility programs. Appl Energy 2017;203:128–41. https://doi.org/10.1016/j.apenergy.2017.06.040.
[39] Hoog JD, Alpcan T, Brazil M, Thomas DA, Mareels I. Optimal charging of electric vehicles taking distribution network constraints into account. IEEE Trans Power

Syst 2014;30(1):365–75. https://doi.org/10.1109/PESGM.2017.8273990.
[40] Madani R, Sojoudi S, Lavaei J. Convex relaxation for optimal power flow problem: mesh networks. IEEE Trans Power Syst 2015;30(1):199–211. https://doi.org/10.1109/TPWRS.2014.2322051.
[41] Wei T, Zhu Q, Yu N. Proactive demand participation of smart buildings in smart grid. IEEE Trans Comp 2016;65(5):1392–406. https://doi.org/10.1109/TC.2015.2495244.
[42] Pudjianto D, Ramsay C, Strbac G. Virtual power plant and system integration of distributed energy resources. IET Renew Power Gener 2007;1(1):10–6. https://doi.org/10.1049/iet-rpg:20060023.
[43] Distribution test feeders - distribution test feeder working group - IEEE PES distribution system analysis subcommittee < https://ewh.ieee.org/soc/pes/dsacom/testfeeders > .
[44] Zhai Q, Guan X, Cui J. Unit commitment with identical units successive subproblem solving method based on Lagrangian relaxation. IEEE Trans Power Syst 2002;17(4):1250–7. https://doi.org/10.1109/TPWRS.2002.805003.
[45] Mirakhorli A, Dong B. Market and behavior driven predictive energy management for residential buildings. Sustain Cities Soc. https://doi.org/10.1016/j.scs.2018.01.030.
[46] Shi J, Yu N, Yao W. Energy efficient building HVAC control algorithm with real-time occupancy prediction. In: International Conference on Sustainability in Energy and Buildings; 2016. p. 1–10. https://doi.org/10.1016/j.egypro.2017.03.028.
[47] Sun B, Luh PB, Jia QS, Jiang Z, Wang F, Song C. Building energy management: Integrated control of active and passive heating, cooling, lighting, shading, and ventilation systems. IEEE Trans Autom Sci Eng 2013;10(3):588–602. https://doi.org/10.1109/TASE.2012.2205567.
[48] Xu Z, Liu S, Hu G, Spanos CJ. Optimal coordination of air conditioning system and personal fans for building energy efficiency improvement. Energy Build 141. https://doi.org/10.1016/j.enbuild.2017.02.051.
[49] Lu N, Nguyen T. Grid friendly™ appliances - load-side solution for congestion management. Transmission and distribution conference and exhibition, 2005/2006 IEEE PES IEEE; 2006. p. 1269–73. https://doi.org/10.1109/TDC.2006.1668693.
[50] Godina R, Rodrigues EM, Matias JC, Catalo JP. Smart electric vehicle charging scheduler for overloading prevention of an industry client power distribution transformer. Appl Energy 2016;178:29–42. https://doi.org/10.1016/j.apenergy.2016.06.019.
[51] Fisher ML. The Lagrangian relaxation method for solving integer programming problems. INFORMS; 2004.
[52] Zhao X, Luh PB, Wang J. Surrogate gradient algorithm for Lagrangian relaxation. J Optim Theory Appl 1999;100(3):699–712. https://doi.org/10.1023/A:1022646725208.
[53] Wang W, Yu N. LMP decomposition with three-phase DCOPF for distribution system. In: 2016 IEEE innovative smart grid technologies - Asia (ISGT-Asia); 2016. p. 1–8. https://doi.org/10.1109/ISGT-Asia.2016.7796352.
[54] Li F, Bo R. DCOPF-based LMP simulation: algorithm, comparison with ACOPF, and sensitivity. IEEE Trans Power Syst 2007;22(4):1475–85. https://doi.org/10.1109/TPWRS.2007.907924.
[55] Löfberg J. YALMIP: a toolbox for modeling and optimization in MATLAB. In: Proceedings of the CACSD conference. Taipei (Taiwan); 2004. https://doi.org/10.1109/CACSD.2004.1393890.
[56] Gurobi I. Optimization, Gurobi optimizer reference manual; 2016 < http://www.gurobi.com > .
[57] Candanedo LM, Feldheim V. Accurate occupancy detection of an office room from light, temperature, humidity and CO_2 measurements using statistical learning models. Energy Build 2016;112:28–39. https://doi.org/10.1016/j.enbuild.2015.11.071.
[58] Dong B. Long-term occupancy data for residential and commercial building. openEI < https://openei.org/datasets/dataset/long-term-occupancy-data-for-residential-and-commercial-building > .
[59] Lu N. An evaluation of the HVAC load potential for providing load balancing service. IEEE Trans Smart Grid 2012;3(3):1263–70. https://doi.org/10.1109/TSG.2012.2183649. 00137.
[60] Schneider KP, Fuller JC, Chassin DP. Multi-state load models for distribution system analysis. IEEE Trans Power Syst 2011;26(4):2425–33. https://doi.org/10.1109/TPWRS.2011.2132154.
[61] Liu R, Dow L, Liu E. A survey of PEV impacts on electric utilities. In: ISGT 2011; 2011. p. 1–8. https://doi.org/10.1109/ISGT.2011.5759171.
[62] Wang W, Yu N, Foggo B, Davis J, Li J. Phase identification in electric power distribution systems by clustering of smart meter data. In: 2016 15th IEEE international conference on machine learning and applications (ICMLA); 2016. p. 259–265. https://doi.org/10.1109/ICMLA.2016.0050.
[63] Wang W, Yu N. Chordal conversion based convex iteration algorithm for three-phase optimal power flow problems. IEEE Trans Power Syst 2018;33(2):1603–13. https://doi.org/10.1109/TPWRS.2017.2735942.