

Online Scheduling of Thermostatically Controlled Loads: A Stepwise Reformulation Approach

Xueyuan Cui, *Member, IEEE*, Liudong Chen, *Graduate Student Member, IEEE*,
Yi Wang, *Senior Member, IEEE*, and Bolun Xu, *Member, IEEE*

Abstract—Thermostatically controlled loads (TCLs) in buildings serve as major flexible resources in power systems, and their online scheduling is critical for providing real-time demand response. However, the inherent coupling arising from inter-zone thermal dynamics and multi-source uncertainties (e.g., model approximation errors and imperfect price forecasting) renders the resulting stochastic optimization problem computationally intractable for real-time applications. This work proposes a stepwise modeling-optimization reformulation approach to simplify the problem and accelerate the online scheduling. In particular, a clustering-based thermal dynamics partitioning strategy is developed to aggregate neighboring zones with strong thermal interactions, thereby decoupling the high-dimensional thermal dynamics into individual single-zone models to facilitate parallel computation. Furthermore, a deterministic transformation method is introduced to convert chance constraints and expected-value objectives into equivalent deterministic forms. Through this stepwise reformulation, the original problem is transformed into a decoupled, deterministic, and single-stage optimization that can be efficiently solved online. Numerical simulations on a 100-zone building prototype demonstrate that the proposed approach significantly enhances computational efficiency while maintaining economic optimality for online TCL scheduling.

Index Terms—Building energy systems, thermostatically controlled loads, thermal dynamics, chance constraints, dynamic programming.

NOMENCLATURE

A. Indices and Sets

i, j	Indices of thermal zones.
c	Index of equivalent clusters.
t	Index of time slots.
k	Index of uncertainty samples.
$\mathcal{Z}$	Set of original thermal zones.
$\mathcal{Z}_c$	Set of original zones belonging to cluster c .

X. Cui and Y. Wang were supported in part by the National Natural Science Foundation of China (T2541025), in part by the Research Grants Council of the Hong Kong SAR (HKU 17201225), and in part by the Research Grants Council of the Hong Kong SAR (CRF-YCRG C1002-23Y). (*Corresponding author: Yi Wang*)

Xueyuan Cui is with the Department of Electrical and Computer Engineering, The University of Hong Kong, Hong Kong SAR, China, with the Shenzhen Institute of Research and Innovation, The University of Hong Kong (Shenzhen), Shenzhen 518057, China, and with the Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA, (email: xcui@seas.harvard.edu)

Yi Wang is with the Department of Electrical and Computer Engineering, The University of Hong Kong, Hong Kong SAR, China, and is also with the Shenzhen Institute of Research and Innovation, The University of Hong Kong (Shenzhen), Shenzhen 518057, China (email: yiwang@eee.hku.hk)

Liudong Chen and Bolun Xu are with the Department of Earth and Environmental Engineering, Columbia University, New York, NY 10025, USA (email: lc3671@columbia.edu, bx2177@columbia.edu).

$\mathcal{C}$	Set of clusters.
$\mathcal{T}$	Set of scheduling periods.
$\mathcal{E}$	Set of edges for inter-zone thermal interactions.
$\mathcal{K}$	Set of sampled uncertainty scenarios.

B. Parameters

C_i	Thermal capacitance of zone i .
R_i	Thermal resistance between zone i and the outdoor environment.
R_{ij}	Thermal resistance between zones i and j .
η_i	Coefficients of the temperature term in zone i .
β_i	Coefficients of the TCL power term in zone i .
α_{ij}	Thermal interaction coefficient from zone j to zone i .
$\bar{\alpha}_{ij}$	Symmetric thermal interaction coefficient between zones i and j .
δ_t	External disturbance from outdoor temperature.
$\delta'_{i,t}$	The i -th zone's internal disturbance from occupants, lighting, and other equipment.
$\bar{\varepsilon}_{i,t}$	Disturbance term in thermal dynamics of zone i .
$\varepsilon_{i,t}, \varepsilon_{c,t}$	Transformed disturbance terms for original zone i and equivalent cluster c .
$s'_{i,t}$	Reference indoor temperature of zone i .
λ_t	Real-time electricity price at time t .
$\rho_{i,t}$	Comfort penalty coefficient of arbitrary zone i at time t .
$\bar{P}_i, P_c$	Power capacity of TCLs in zone i and equivalent cluster c .
$\underline{S}_{i,t}, \bar{S}_{i,t}$	Lower/upper indoor-temperature limits of zone i .
$\underline{S}_{c,t}, \bar{S}_{c,t}$	Lower/upper equivalent-temperature limits of cluster c .
γ	Risk level of the chance constraint.
$\mu_{i,t}, \sigma_{i,t}$	Mean and standard deviation of model error in zone i .
$\underline{\xi}_{c,t}, \bar{\xi}_{c,t}$	Lower/upper uncertainty bounds used in deterministic constraints.
ω	Weight coefficient in the thermal dynamics partitioning objective.
$\underline{M}_c, \bar{M}_c$	Minimum/maximum number of zones in cluster c .
F_t, f_t	Cumulative distribution and probability density functions of real-time price at time t .
$U_{i,t}$	Comfort-violation penalty function.
$V_{t= \mathcal{T} }$	Terminal value function.

C. Variables

$\bar{s}_{i,t}$	Indoor temperature of original zone i before reference transformation.
$s_{i,t}$	Temperature deviation of original zone i from its reference value.
$\tilde{s}_{i,t}$	Actual temperature deviation of original zone i under model error.
$s_{c,t}$	Equivalent temperature deviation of cluster c .
$p_{i,t}$	TCL power output in original zone i .
$p_{c,t}$	Equivalent TCL power output in cluster c .
$\xi_{i,t}, \xi_{c,t}$	Model errors of original zone i and equivalent cluster c .
C	Total scheduling cost.
$b_{i,c}$	Binary variable indicating whether zone i belongs to cluster c .
$r_{i,j,c}$	Binary variable indicating whether edge (i, j) is included in cluster c .
$e_{i,j}$	Binary variable indicating whether edge (i, j) is within any cluster.
V_t, Q_t	Value function and minimized cost at time t .
v_t, q_t	Derivatives of $V_t(\cdot)$ and $Q_t(\cdot)$.
θ_t	Dual variable associated with equivalent thermal dynamics.
w_t	Sum of the derivatives of value and comfort-penalty functions.
$p_{i,t}^*, s_{i,t}^*$	Optimal TCL output and temperature state of zone i .

I. INTRODUCTION

Thermostatically controlled loads (TCLs) account for approximately 40% of total energy consumption in buildings [1]. TCLs also serve as critical flexible resources for demand response in power grids by leveraging thermal inertia in building zones to regulate indoor temperature [2]. Online scheduling of TCLs is an effective way to harness the power consumption flexibility while satisfying human comfort preferences in real time.

However, online scheduling of TCLs in buildings remains a significant challenge due to thermal interactions and uncertainties. In buildings, different zones represent distinct spaces equipped with a thermostat-controlled Heating, Ventilation, and Air Conditioning (HVAC) system to maintain thermal comfort [3]. The multi-zone scheduling requires modeling thermal dynamics to capture inter-zonal interactions. Such dynamics are highly nonlinear and are coupled with neighboring zones through heat transfer, and the modeling dimension of thermal dynamics increases with the number of zones, further complicating the optimization-based TCL scheduling [4]. Besides, uncertainties in the TCL scheduling problem could arise from approximation errors in modeling building thermal dynamics and from imperfect forecasting of external electricity prices. Such multi-source uncertainties could amplify deviations in scheduling decisions, leading to higher economic costs. To sum up, inter-zone thermal interactions and multi-source uncertainties jointly make the original online TCL scheduling problem coupled, stochastic, and multi-stage for solving.

Given the challenges, current research primarily focuses on advanced modeling or optimization algorithms to improve decision accuracy and computational efficiency of on-line scheduling. Modeling methods rely on dimensionality reduction for the original high-dimensional thermal dynamics. Specifically, order reduction [5], [6] and model aggregation [7] methods effectively simplify high-dimensional optimization with reduced-order models. However, thermal interaction exists in these derived models, which could cause error propagation in scheduling results across zones. Besides, there remains a lack of clarity on how to recover scheduling decisions from the reduced-order models.

Recent advances in optimization algorithms have shown promise in addressing these challenges. Distributed optimization methods can effectively decouple thermal dynamics across zones to support local scheduling; however, ignoring thermal interactions can lead to suboptimal decisions for each zone. Model predictive control (MPC) has emerged as a prominent method for addressing price and model uncertainties in TCL scheduling, which involves solving a finite-horizon optimization problem using forecasted data and implementing the first-stage TCL scheduling decision [8]. However, MPC is computationally intensive because it requires repeatedly solving optimization problems in a receding-horizon framework. Alternatively, reinforcement learning (RL) [9] could generate precise scheduling decisions by training on large-scale temperature control scenarios. However, RL requires discretizing actions, states, and prices to formulate Markov decision processes (MDPs), and it can encounter the ‘‘curse of dimensionality’’: computational time grows exponentially with the number of decision stages and scenarios.

Thus, *how can we solve the online TCL scheduling problem efficiently and accurately, with coupled thermal interactions and multi-source uncertainties?* Currently, modeling-based methods can reduce the dimension of multi-zone thermal dynamics, but the resulting reduced models may still retain residual thermal coupling and may not provide a direct way to recover feasible zone-level scheduling decisions. Optimization-based methods (e.g., distributed optimization, MPC, and RL) can improve scalability or uncertainty handling, but they could rely on simplified decoupled models, repeated online solution of multi-stage stochastic problems, or discretized state-action spaces. Therefore, a key challenge is how to simultaneously handle thermal coupling and multi-source uncertainties while preserving the computational efficiency required for online applications. In this work, we propose an integrated modeling-optimization framework to balance the accuracy and complexity of high-dimensional uncertain optimization problems. The proposed stepwise reformulation approach sequentially decouples, detrends, and single-stages the original coupled, stochastic, and multi-stage problem in the offline stage, thereby accelerating the solution process for online scheduling without sacrificing solution accuracy relative to the original problem. The key contributions are as follows:

- 1) We formulate a stepwise reformulation approach for online TCL scheduling in multi-zone buildings. In this approach, we combine modeling and optimization algorithms sequentially to convert the original coupled,

stochastic, and multi-stage optimization problem into a decoupled, deterministic, and single-stage one. Benefiting from parallel solving across zones with well-developed solvers for deterministic optimization, the reformulated problem for online scheduling can effectively improve solving efficiency while maintaining economic costs and thermal comfort.

- 2) We develop a clustering-based thermal-dynamics partitioning strategy to mitigate thermal interactions across neighboring zones. We first transfer building structures into a thermal graph, which is a graph-based representation of multi-zone thermal dynamics where each zone is treated as a node and each inter-zone thermal interaction is represented as an edge. Then, we develop a stochastic optimization problem to minimize the cross-cluster interaction weights. Neighboring zones with stronger interaction are aggregated into a cluster, transforming the original coupled thermal dynamics into a decoupled model with reduced aggregated zones.
- 3) We propose a two-stage deterministic transformation method for handling multi-source uncertainties. We first convert chance constraints, which impose probabilistic comfort requirements under model errors, into equivalent deterministic comfort bounds. Furthermore, we derive an analytical stochastic dynamic programming (ASDP) to transform the expected-value objective arising from probabilistic forecasting of real-time prices (RTP) into an equivalent, deterministic function. In dynamic programming, the deterministic problem can be solved in a single stage for each scheduling time.

The remainder of this work is organized as follows: Section II presents the literature review. Section III formulates the research problem for online TCL scheduling, and Section IV introduces the proposed solution approach. Section V conducts case studies. Conclusions and future works are reached in Section VI.

II. LITERATURE REVIEW

A common research direction focuses on reducing computational complexity by using advanced modeling techniques to derive lower-dimensional equivalent representations. For multi-zone systems, high-fidelity reduced-order models using controllability–observability analysis and balanced truncation enable tractable MPC-based TCL scheduling [5]. Lumped-parameter representations further simplify thermal dynamics and facilitate online optimization within MPC frameworks [7]. Virtual energy storage systems are adopted to capture the impact of multiple temperature zones due to TCLs’ switching and occupants’ activity [10], [11]. In parallel with physics-based approaches, data-driven methods have been widely explored to construct surrogate thermal models. Learning-based techniques have been used to extract low-dimensional latent representations that can be embedded into optimization frameworks for TCL scheduling [6]. To cope with limited measurement availability, state estimation methods, such as the Unscented Kalman Filter [12] and Vector Fitting [13], have been combined with optimization to support multi-objective

scheduling decisions. More recently, neural-network-based models have been trained to approximate thermal dynamics, with deep learning surrogates further accelerating the solution of TCL scheduling problems [14]. Yet, despite the diversity of physics-driven and data-driven model reduction strategies, most existing approaches retain coupled thermal interactions in the reduced models. Such residual coupling can induce error propagation across zones, ultimately limiting scheduling performance in large-scale multi-zone building systems.

Related work has also investigated advanced optimization algorithms for coupled and uncertain TCL scheduling. Distributed optimization frameworks are widely used to decompose large-scale multi-zone problems into tractable zone-level subproblems, often within hierarchical architectures that coordinate building-level objectives and local control actions [15]. Multi-agent distributed control further enables real-time operation by allowing each zone to perform local optimization while coordinating with neighboring zones, with both cooperative and non-cooperative formulations explored to enhance scalability [16], [17]. The bilevel optimization reformulation strategy addresses over-conservatism and computational intractability in the traditional distributionally robust chance-constrained TCL scheduling problem [18]. MPC is another prevalent approach for managing forecasting uncertainties in TCL scheduling. Hierarchical MPC schemes separate long-term solving from real-time control, typically using coarse models at higher levels and refining decisions at lower levels [19]. To address uncertainty propagation across zones, distributionally robust chance-constrained formulations have been developed, with inner-approximation methods employed to improve computational tractability [20]. RL has also gained attention for stochastic TCL scheduling by learning control policies directly from large-scale scenarios. Hybrid evolutionary–RL frameworks and multi-agent MDP-based formulations, including federated learning mechanisms, have been proposed to support scalable HVAC control [21]–[23]. However, RL methods typically require discretized state, action, and price spaces, leading to exponential growth in computational complexity as the number of decision stages and scenarios increases.

The conclusion of the literature review for multi-zone TCLs scheduling is presented in Table I. In light of the above context, existing research typically focuses on advanced modeling or optimization techniques to tackle the coupled, stochastic, and multi-stage scheduling. However, two research gaps remain: 1) the order-reduced model could be intractable in the optimization problem to recover decisions of original variables; 2) stochastic optimization could be infeasible to balance the requirements of solution accuracy and computational efficiency for online applications.

III. PROBLEM FORMULATION

A. Multi-zone Thermal Dynamics

In TCL scheduling, thermal dynamics represent the indoor temperature change caused by historical temperature, power output of TCL, and other disturbances (e.g., outdoor temperature and internal heat sources). A notable feature of multi-

Table I
COMPARISON OF MULTI-ZONE TCL SCHEDULING LITERATURE

Ref.	Focus	Zones	Uncertainty	Stages
[5], [7], [10], [11]	Model	Coupled	–	Multi
[6], [12], [13]	Model	Coupled	–	Multi
[14]	Model	Coupled	Stoch.	Multi
[15], [16]	Opt.	Decoupled	–	Multi
[17]–[19]	Opt.	Decoupled	Stoch.	Multi
[20]	Opt.	Coupled	Stoch.	Multi
[21]–[23]	Opt.	Coupled	Stoch.	Single
Proposed	Model+Opt.	Decoupled	Det.	Single

Stoch.: Stochastic, Det.: Deterministic, Opt.: Optimization

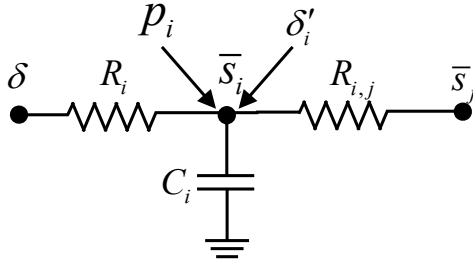


Figure 1. Resistance-capacitance based thermal dynamics model.

zone buildings is that heat exchange may happen from higher-temperature to lower-temperature zones. Such thermal interaction would influence indoor temperature when scheduling TCL outputs of neighboring zones. As shown in Fig. 1, taking the arbitrary i -th zone as an example, the thermal dynamics are modeled using the first-order differential equation as

$$C_i \frac{d\bar{s}_i}{dt} = \frac{\delta - \bar{s}_i}{R_i} + \sum_{j \in \mathcal{Z} \setminus \{i\}} \frac{\bar{s}_j - \bar{s}_i}{R_{ij}} + p_i + \delta'_i, \forall i \in \mathcal{Z}, \quad (1)$$

In this model, δ, δ'_i denote the external disturbance (outdoor temperature) and the i -th zone's internal disturbance (heat gains from occupants, lighting, and other equipment), and p_i denotes TCL power output in the i -th zone (heating power in this work). Both $\bar{s}$ and δ are regarded as thermal nodes. As indoor temperature is the state variable, capacitance C_i is required to represent the effect of thermal inertia. For the disturbance δ , only resistance R_i is needed to represent the thermal conduction process from the ambient to the indoor thermal environment.

Using the Euler method, the discrete-form thermal dynamics can be written as

$$\bar{s}_{i,t} = \eta_i \bar{s}_{i,t-1} + \sum_{j \in \mathcal{Z} \setminus \{i\}} \alpha_{ij} \bar{s}_{j,t-1} + \beta_i p_{i,t} + \bar{\varepsilon}_{i,t} \quad (2a)$$

$$\eta_i = 1 - \frac{1}{C_i R_i} - \sum_{j \in \mathcal{Z} \setminus \{i\}} \frac{1}{C_i R_{ij}} \quad (2b)$$

$$\beta_i = \frac{1}{C_i}, \alpha_{ij} = \frac{1}{C_i R_{ij}} \quad (2c)$$

$$\bar{\varepsilon}_{i,t} = \frac{\delta_t}{C_i R_i} + \frac{\delta'_{i,t}}{C_i} \quad (2d)$$

Generally, temperature regulation should follow a reference curve, which we denote as $s'_{i,t}$. Then we denote $s_{i,t} = \bar{s}_{i,t} - s'_{i,t}$ as the violation degree of indoor temperature compared with the reference value. With the definition of $s_{i,t}$, we transform the original model into

$$s_{i,t} = \eta_i s_{i,t-1} + \sum_{j \in \mathcal{Z} \setminus \{i\}} \alpha_{ij} s_{j,t-1} + \beta_i p_{i,t} + \varepsilon_{i,t} \quad (3a)$$

$$\varepsilon_{i,t} = \bar{\varepsilon}_{i,t} - s'_{i,t} + \eta_i s'_{i,t-1} + \sum_{j \in \mathcal{Z} \setminus \{i\}} \alpha_{ij} s'_{j,t-1} \quad (3b)$$

Furthermore, the first-order equation has unavoidable errors in representing actual thermal dynamics. The errors could exist in the identified model parameters R, C and the forecasting of disturbance values δ, δ' . To enhance the model's tractability, we denote the actual temperature of the i -th zone as $\tilde{s}_{i,t}$, which satisfies:

$$\tilde{s}_{i,t} = s_{i,t} + \xi_{i,t} \quad (4a)$$

$$\xi_{i,t} \in \mathcal{N}(\mu_{i,t}, \sigma_{i,t}) \quad (4b)$$

Without loss of generality, we assume the uncertain model errors follow the normal distribution [24], with average value $\mu_{i,t}$ and standard deviation value $\sigma_{i,t}$ at each time t .

B. Chance-constrained Stochastic Optimization

The multi-zone temperature scheduling problem aims to minimize the economic cost for electricity purchase and the penalty cost of thermal comfort violation for the whole period. The total cost function can be formulated as

$$C = \sum_{i \in \mathcal{Z}} \sum_{t \in \mathcal{T}} (\lambda_t \cdot p_{i,t} + U_{i,t}(s_{i,t})) \quad (5)$$

where $U_{i,t}(s_{i,t})$ is the utility function that measures the penalty of violating the reference temperature, and it is typically represented by a quadratic function as:

$$U_{i,t} = \rho_{i,t} \cdot (\bar{s}_{i,t} - s'_{i,t})^2 = \rho_{i,t} \cdot s_{i,t}^2 \quad (6)$$

where $\rho_{i,t}$ denotes the comfort penalty coefficient of arbitrary zone i at time t to measure the sensitivity of discomfort.

Besides, for each zone at each time, the problem is subject to operational limits of TCLs and also the comfort limits of indoor temperature, as

$$0 \leq p_{i,t} \leq P_i, \quad (7a)$$

$$\underline{s}_{i,t} \leq \tilde{s}_{i,t} \leq \bar{s}_{i,t} \quad (7b)$$

The addition of hard constraints (7b) on temperature is helpful to avoid the sudden loss of indoor comfort due to real-time price spikes.

Due to the uncertain model errors ξ , the predicted temperature could deviate from the actual one and violate the hard constraint. Thus, we relax (7b) into chance constraints, which require the indoor temperature to remain within the comfort bounds with a prescribed confidence level rather than for all possible uncertainty realizations, as

$$\Pr(\underline{s}_{i,t} \leq \tilde{s}_{i,t} \leq \bar{s}_{i,t}) \geq 1 - \gamma \quad (8)$$

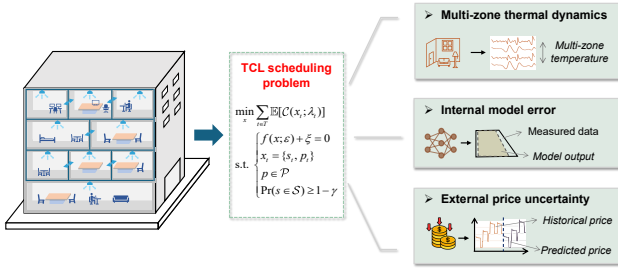


Figure 2. The original TCL scheduling problem.

where $1 - \gamma$ denotes the required confidence level, and γ represents the allowable probability of comfort-bound violation caused by model errors.

The overall stochastic optimization is formulated as

$$\min_{p_{i,t}} \mathbb{E}[C] + V_{t=|T|}(s), \quad (9a)$$

$$\begin{aligned} \text{s.t. } & (3), (4), (7a), (8), \\ & \forall i \in \mathcal{Z}, \forall t \in \mathcal{T} \end{aligned} \quad (9b)$$

where we denote $V_{t=|T|}$ as the value function of the indoor temperature at the final time stage. In this optimization problem, the effects of occupancy are implicitly captured through the internal heat source disturbance term δ' in the thermal dynamics model and zone-dependent comfort bounds $\underline{S}_{i,t}, \bar{S}_{i,t}$. The internal heat source disturbances include occupancy-related heat gains, while different comfort temperature ranges are assigned to different zones by considering the heterogeneous and time-varying comfort preferences of occupants.

Discussion: As shown in Fig. 2, the problem (9) needs to be solved with multi-zone thermal interaction, chance constraints due to internal model errors, and the expected objective function subject to non-anticipatory decision making as the real-time price λ_t is announced sequentially intraday. When applying (9) for online TCL scheduling, the highly coupled, stochastic, and multi-stage nature of the problem may lead to long computational times and biased decisions due to error propagation across zones. Neither the solving efficiency nor the solution feasibility could be satisfied for online scheduling.

IV. SOLUTION APPROACH

In this section, we propose a step-wise approach to reformulate the original problem for online scheduling implementation. An overview of the proposed method is presented in Fig. 3. Specifically, we first decouple the multi-zone thermal dynamics based on the clustering-based thermal dynamics partitioning method, in which the influence of model error on partitioning results is considered. Furthermore, we follow a two-stage deterministic transformation method to convert the chance constraints and expected objective function into deterministic ones. Finally, we obtain an approximated decoupled, deterministic, and single-stage problem for online implementation. By leveraging parallel calculation of each zone and the deterministic transformation, the computational time and the solution deviations could be alleviated effectively.

A. Clustering-based Thermal Dynamics Partitioning

In multi-zone buildings, thermal zones are typically divided based on building orientations, room functions, load characteristics, and the configuration of air-conditioning equipment. Here, we propose to divide or aggregate these zones, aiming to decouple thermal dynamics across all zones. To that end, a clustering-based zone partitioning method is proposed to classify zones into different clusters, which is then reformulated into a stochastic optimization problem to minimize the total thermal interaction among different clusters.

1) *Thermal interaction and graph:* The relative magnitude of thermal interaction between zones i and j are represented as α_{ij} and α_{ji} in the model (3a). It can be seen that the α_{ij} (α_{ji}) is jointly determined by the thermal capacitance C_i (C_j) and the thermal resistance R_{ij} (R_{ji}) in (1), where we have $C_i \neq C_j$ and $R_{ij} = R_{ji}$. The different thermal capacitance of zones causes the asymmetric interaction between α_{ij} and α_{ji} . To that end, we take the average value as the final interaction coefficient as:

$$\bar{\alpha}_{ij} = (\alpha_{ij} + \alpha_{ji})/2 \quad (10a)$$

$$\forall i \in \mathcal{Z}, \quad \forall (i, j) \in \mathcal{E} \quad (10b)$$

Furthermore, treating each zone temperature as a node and inter-zone thermal interactions as edges yields a thermal graph representation of the multi-zone dynamics. In this graph, the node set denotes the thermal zones, while the edge set denotes the physical heat-transfer links between neighboring zones. This representation allows the strength of inter-zone thermal coupling to be quantified and used for clustering. Given that condition, we can denote the thermal graph with the sets of nodes $\mathcal{Z}$ and edges $\mathcal{E}$, respectively.

2) *Stochastic optimization for partitioning:* To further divide the original graph to minimize the overall thermal interactions, we propose a stochastic optimization method for optimal partitioning [25] to mitigate the impacts of model error ξ . The major principle is to aggregate the neighboring zones with larger interactions into the same cluster, while dividing zones with smaller interactions into different clusters. We denote the set of clusters as $\mathcal{C}$, and we denote the binary variable $b_{i,c} \in \{0, 1\}$ indicating whether the i -th zone is in the c -th cluster. Then, we denote the binary variable $r_{i,j,c}$ that represents whether the edge (i, j) is in the c -th cluster, and the binary variable $e_{i,j}$ that represents whether the edge (i, j) is in any clusters.

The objective of thermal dynamics partitioning is to minimize the thermal interaction between zone clusters while maximizing the thermal interaction within each zone cluster. Considering the model error on the parameter α_{ij} , the objective function is formulated as:

$$\min_{e_{i,j}} \sum_{(i,j) \in \mathcal{E}} (1 - e_{i,j}) \cdot \mathbb{E}[\bar{\alpha}_{ij}] - \omega \sum_{c \in \mathcal{C}} \sum_{(i,j) \in \mathcal{E}} r_{i,j,c} \cdot \mathbb{E}[\bar{\alpha}_{ij}] \quad (11)$$

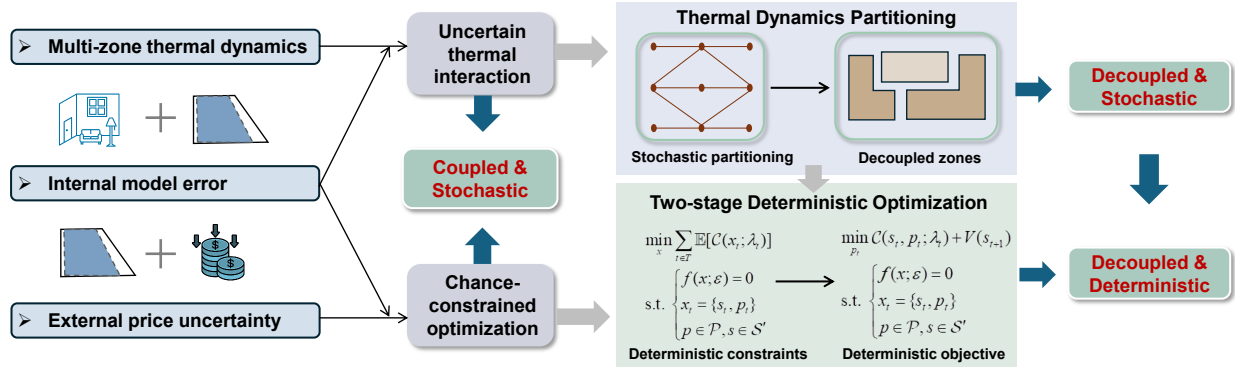


Figure 3. Overall route of the proposed solution approach.

This clustering problem is subject to the following constraints:

$$\sum_{c \in \mathcal{C}} b_{i,c} = 1, \quad \forall i \in \mathcal{Z} \quad (12a)$$

$$\sum_{i \in \mathcal{Z}} b_{i,c} \geq 1, \quad \forall c \in \mathcal{C} \quad (12b)$$

$$r_{i,j,c} \leq b_{i,c}, \quad \forall (i,j) \in \mathcal{E}, \forall c \in \mathcal{C} \quad (12c)$$

$$r_{i,j,c} \leq b_{j,c}, \quad \forall (i,j) \in \mathcal{E}, \forall c \in \mathcal{C} \quad (12d)$$

$$b_{i,c} + b_{j,c} \leq r_{i,j,c} + 1, \quad \forall (i,j) \in \mathcal{E}, \forall c \in \mathcal{C} \quad (12e)$$

$$e_{i,j} = \sum_{c \in \mathcal{C}} r_{i,j,c}, \quad \forall (i,j) \in \mathcal{E} \quad (12f)$$

$$\underline{M}_c \leq \sum_{i \in \mathcal{Z}} b_{i,c} \leq \overline{M}_c \quad c \in \mathcal{C} \quad (12g)$$

where (12a) indicates that each zone is included in only one cluster; (12b) indicates that each cluster includes at least one zone; (12c)-(12d) indicate that both of the nodes are included into a cluster when the edge is included; (12e) indicates that the edge is included in a cluster when both of the nodes are included; (12f) is used to know if the edge is included in any cluster. Besides these constraints on the connection relationship of zones in clusters, we further denote (12g) to limit the size of each cluster within thresholds $\underline{M}_c$ and $\overline{M}_c$.

This stochastic mixed integer linear programming (MILP) can be solved with the expected value of $\bar{\alpha}$ as input, which can be evaluated by Monte-Carlo simulation when identifying the original thermal dynamics model.

3) *Model aggregation*: After zone partitioning, there could be more than one zone in each cluster. We denote the interaction on edges with $e_{i,j} = 0$ (i.e., crossing edges between clusters) as 0 to divide each cluster. For zones in each cluster, we aggregate them as an equivalent single zone. Given the original form of thermal dynamics in (2), this aggregation process could be formulated as:

$$\begin{aligned} \sum_{i \in \mathcal{Z}_c} C_i s_{i,t} &= \sum_{i \in \mathcal{Z}_c} C_i s_{i,t-1} - \sum_{i \in \mathcal{Z}_c} \frac{s_{i,t-1}}{R_i} + \sum_{i \in \mathcal{Z}_c} p_{i,t} \\ &+ \sum_{i \in \mathcal{Z}_c} \sum_{j \in \mathcal{Z}_c \setminus \{i\}} \frac{s_{j,t-1} - s_{i,t-1}}{R_{ij}} + \sum_{i \in \mathcal{Z}_c} C_i \varepsilon_{i,t} \end{aligned} \quad (13)$$

Given the above formulation, the thermal interaction becomes zero because we have $R_{ij} = R_{ji}$. In this way, we obtain an equivalent zone for each cluster. These zones are without any thermal interaction, i.e.,

$$s_{c,t} = \eta_c s_{c,t-1} + \beta_c p_{c,t} + \varepsilon_{c,t}, \quad \forall c \in \mathcal{C} \quad (14a)$$

$$s_{c,t} = \frac{\sum_{i \in \mathcal{Z}_c} C_i s_{i,t}}{\sum_{i \in \mathcal{Z}_c} C_i}, \quad p_{c,t} = \sum_{i \in \mathcal{Z}_c} p_{i,t} \quad (14b)$$

$$\varepsilon_{c,t} = \frac{\sum_{i \in \mathcal{Z}_c} C_i \varepsilon_{i,t}}{\sum_{i \in \mathcal{Z}_c} C_i} \quad (14c)$$

$$\eta_c \approx 1 - \frac{\sum_{i \in \mathcal{Z}_c} \frac{1}{R_i}}{\sum_{i \in \mathcal{Z}_c} C_i} \quad (14d)$$

$$\beta_c = \frac{1}{\sum_{i \in \mathcal{Z}_c} C_i} \quad (14e)$$

At the same time, the limits for equivalent temperature and power can be converted as follows:

$$\bar{S}_{c,t} = \frac{\sum_{i \in \mathcal{Z}_c} C_i \bar{S}_{i,t}}{\sum_{i \in \mathcal{Z}_c} C_i}, \quad \forall c \in \mathcal{C} \quad (15a)$$

$$\underline{S}_{c,t} = \frac{\sum_{i \in \mathcal{Z}_c} C_i \underline{S}_{i,t}}{\sum_{i \in \mathcal{Z}_c} C_i}, \quad \forall c \in \mathcal{C} \quad (15b)$$

$$P_c = \sum_{i \in \mathcal{Z}_c} P_i \quad (15c)$$

Discussion: Note that the equivalent cluster-level comfort bounds $\bar{S}_c, \underline{S}_c$ may cause zone-level comfort violations due to different $\bar{S}_i, \underline{S}_i$ across zones. Specifically, we make the following solutions to mitigate the remaining error: first, the equivalent cluster comfort bounds are generated conservatively from the original zone-level bounds in (15), so the cluster-level feasible region does not relax the original comfort requirements; second, aggregation errors are considered in the deterministic reformulation of the chance constraints (16), which tightens the feasible temperature range according to the prescribed risk level; third, during online implementation, the actual zone-level temperatures are measured after control execution and re-aggregated as the initial state for the next scheduling step (i.e., Lines 12-14 in Algorithm 2), which prevents aggregation errors from accumulating over time.

Algorithm 1: Clustering-based Thermal Dynamics Partitioning

Input: $\mathcal{Z}; \mathcal{E}; \alpha_{ij}, \alpha_{ji}; \omega; \underline{M}_c, \overline{M}_c$.

Output: $\mathcal{C}; \mathcal{Z}_c (\forall c \in \mathcal{C}); \underline{S}_{c,t}, \underline{s}_{c,t}, P_c (\forall c \in \mathcal{C})$.

- 1 **Thermal graph construction:**
- 2 Build a thermal graph with each zone $i \in \mathcal{Z}$ as a node and each thermal interaction $(i, j) \in \mathcal{E}$ as an edge;
- 3 For each edge $(i, j) \in \mathcal{E}$, calculate $\bar{\alpha}_{ij}$ based on (10);
- 4 Evaluate $\mathbb{E}[\bar{\alpha}_{ij}]$ considering the distribution of model errors;
- 5 **Stochastic partitioning:**
- 6 Formulate the thermal dynamics partitioning problem with objective function (11) and constraints in (12);
- 7 Solve the resulting stochastic MILP and obtain the optimal clustering result to construct $\mathcal{Z}_c = \{i \in \mathcal{Z} \mid b_{i,c} = 1\}$;
- 8 **Model aggregation:**
- 9 For edges with $e_{i,j} = 0$, ignore the corresponding inter-cluster thermal interactions;
- 10 **for each cluster $c \in \mathcal{C}$ do**
- 11 Aggregate all original zones in $\mathcal{Z}_c$ into one equivalent thermal zone;
- 12 Derive the independent equivalent thermal dynamics of cluster c based on (14);
- 13 Convert the original temperature and TCL power limits into equivalent cluster-level limits by (15).

The overall process of the proposed clustering method is presented in Algorithm 1. The critical steps include: 1) constructing a thermal graph to characterize the inter-zone thermal interactions and quantifying the expected interaction weights under model errors; 2) solving the stochastic MILP to obtain the optimal zone partitioning; and 3) aggregating the original zones within each cluster into an independent equivalent thermal model with transformed temperature and power limits for subsequent deterministic optimization.

B. Two-stage Deterministic Transformation

After thermal dynamics partitioning, the task becomes solving the chance-constrained stochastic optimization in (9) independently for every aggregated single-zone model. The major route of the proposed two-stage transformation is presented in Fig. 4. Specifically, we first transform the chance constraints into deterministic styles by quantifying the comfort bounds given the distribution of model errors. Then, we employ the proposed analytical stochastic dynamic programming (ASDP) method to transform the expected objective into a deterministic one. ASDP is developed as an analytical variant of stochastic dynamic programming: instead of solving a multi-stage stochastic optimization problem online, it quantifies the value function of each time step offline under deterministic linear constraints, so that the online scheduling problem can be solved as a single-stage deterministic optimization problem. After the two-stage transformation, the whole problem becomes deterministic and single-stage, which is efficient for online solving.

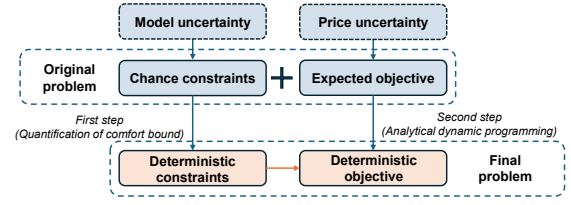


Figure 4. Route of two-stage deterministic transformation.

1) *Deterministic Constraints Transformation:* To handle the chance constraints caused by model errors ξ , we employ the scenario approach to transform them into a tractable form, which can then be reformulated with deterministic bounds.

Proposition 1 (Quantification of comfort bounds). The chance constraints in (8) can be conservatively approximated by the following deterministic linear intervals:

$$\underline{S}_{c,t} - \underline{\xi}_{c,t} \leq s_{c,t} \leq \overline{S}_{c,t} - \bar{\xi}_{c,t}, \quad \forall c \in \mathcal{C}, \forall t \in \mathcal{T} \quad (16)$$

where the bounds $\underline{\xi}_{c,t}$ and $\bar{\xi}_{c,t}$ are quantified based on the initial error distribution and the prediction horizon:

$$\bar{\xi}_{c,t} = \mu_{c,0} + t \cdot \sigma_{c,0} \cdot F_{\mathcal{N}}^{-1}(1 - \gamma) \quad (17a)$$

$$\underline{\xi}_{c,t} = \mu_{c,0} - t \cdot \sigma_{c,0} \cdot F_{\mathcal{N}}^{-1}(1 - \gamma) \quad (17b)$$

where $\mu_{c,0}$ and $\sigma_{c,0}$ represent the mean and standard deviation of the error samples at the initial time step $t = 0$, and $F_{\mathcal{N}}^{-1}(\cdot)$ denotes the inverse cumulative distribution function of (4b).

Proof. The derivation of Proposition 1 relies on the scenario approach for robust optimization and statistical error propagation. Specifically, by introducing Lemmas 3.1-3.2 in [26], we denote $\Pr[g(p, \xi) \leq 0] \geq 1 - \gamma$ as a chance constraint with $g(p, \xi)$ as a function convex in p for any ξ . Then we have

$$\Pr[\Pr[g(p, \xi) \leq 0] \geq 1 - \gamma] \geq 1 - \zeta \quad (18)$$

when the deterministic constraint $g(p, \xi(k)) \leq 0$ holds for all samples $k \in \mathcal{K}$. The size of $\mathcal{K}$ is a sufficiently large number of independent observations, which is bounded by:

$$|\mathcal{K}| \geq \frac{1 + |\mathcal{T}| + \ln(1/\zeta) + \sqrt{2(|\mathcal{T}| + 1) \ln(1/\zeta)}}{\gamma} \quad (19)$$

Furthermore, the probabilistic constraints in (8) can be replaced by a finite set of deterministic constraints across the sample set $\mathcal{K}$:

$$\underline{S}_{c,t} \leq s_{c,t} + \xi_{c,t}(k) \leq \overline{S}_{c,t}, \quad \forall k \in \mathcal{K} \quad (20)$$

To ensure (20) holds for all k , $s_{c,t}$ needs to satisfy the most restrictive cases, i.e.,:

$$s_{c,t} \geq \underline{S}_{c,t} - \min_{k \in \mathcal{K}} \{\xi_{c,t}(k)\}, s_{c,t} \leq \overline{S}_{c,t} - \max_{k \in \mathcal{K}} \{\xi_{c,t}(k)\} \quad (21)$$

By defining $\underline{\xi}_{c,t} = \min_{k \in \mathcal{K}} \{\xi_{c,t}(k)\}$ and $\bar{\xi}_{c,t} = \max_{k \in \mathcal{K}} \{\xi_{c,t}(k)\}$, we obtain the form in (16). Finally, assuming the prediction error $\xi_{c,t}$ follows a normal distribution that scales with time, the extreme values $\underline{\xi}_{c,t}$ and $\bar{\xi}_{c,t}$ are estimated using the $1 - \gamma$ quantiles of the distribution as shown in (17), which completes the proof. $\square$

Given the bounded comfort range, the chance-constrained stochastic optimization problem of (9) after the thermal dynamics partitioning becomes

$$\min_{p_{c,t}} \mathbb{E}[C] + V_{t=|\mathcal{T}|}(s), \quad (22a)$$

$$\text{s.t. (14a), (16),} \quad (22b)$$

$$0 \leq p_{c,t} \leq P_c, \quad (22c)$$

$$\forall c \in \mathcal{C}, \forall t \in \mathcal{T}$$

2) *Deterministic Objective Transformation*: In this subsection, we further mitigate the issue of probabilistic forecasting of real-time prices for deterministic solving of (22). With dynamic programming, we introduce how to solve the non-anticipatory constraint in the equivalent single-zone optimization by recursively solving a single-stage problem backward:

$$Q_{t-1}(s_{t-1}) = \min_{p_t} \{ \lambda_t p_t + U_t + V_t(s_t) | \text{s.t. (22b), (22c)} \} \quad (23)$$

where Q_{t-1} is the minimized operating cost given s_{t-1} , and the value function satisfies that $V_t(s_t) = \mathbb{E}[Q_t(s_t)]$. Similarly, denoting $v_t = \dot{V}_t$ and $q_t = \dot{Q}_t$, we have $v_t(s_t) = \mathbb{E}[q_t(s_t)]$. The index c is avoided for simplification.

Remark 1 (*Convexity of comfort penalty and value function*). It is reasonable to denote the comfort penalty $U_t(s_t)$ as a continuous and convex function, satisfying $U_t(s_t) \geq 0$, $U_t(0) = 0$. Its derivative $u_t = \dot{U}_t$ is monotonic increasing [27]. Meanwhile, the value function V_t is convex and differentiable, and its derivative v_t is also monotonic increasing.

Proposition 2 (*Analytical value function*). The value function $V_t(s_t)$ can be derived analytically for deterministic solving. Specifically, $v_t(s)$ can be written as

$$\begin{aligned} v_{t-1}(s) &= \int_{-\infty}^{\infty} x \cdot \Pr[q_{t-1}(s) = x] dx \\ &= \eta w_t(\eta s + \beta P + \varepsilon_t) F_t(-\beta w_t(\eta s + \beta P + \varepsilon_t)) \\ &\quad - \frac{\eta}{\beta} \cdot \int_{-\beta w_t(\eta s + \beta P + \varepsilon_t)}^{-\beta w_t(\eta s + \varepsilon_t)} x f_t(x) dx \\ &\quad + \eta w_t(\eta s + \varepsilon_t) [1 - F_t(-\beta w_t(\eta s + \varepsilon_t))]. \end{aligned} \quad (24)$$

where $F_t(\cdot)$, $f_t(\cdot)$ are the cumulative distribution function and the probability density function of real-time price at time t .

When solving problem (23), $v_t(s)$ could be discretized by setting s as equally spaced samples. This discretization of $v_t(s)$ corresponds to approximating $V_t(s)$ using piecewise linear functions, and (23) can be solved as a mixed integer linear programming (MILP) problem.

Proof. Following the derivations in the prior research [28], we first apply dual decomposition to obtain the iterative solution of decision variables, which can be utilized to derive the probability density of $q_{t-1}(s_{t-1})$, and then it can be used to quantify v_{t-1} due to the relationship $v_t(s_t) = \mathbb{E}[q_t(s_t)]$. The complete proof is given in Appendix A. $\square$

C. Overall Algorithm

After the stepwise reformulation, the original multi-stage, coupled, and stochastic problem becomes a single-stage, de-

Algorithm 2: Online Scheduling of Multi-zone TCLs

Input: $f(\lambda)$: distribution of forecasting prices; (μ, σ) : average standard deviation of model errors;
 $V_{t=|\mathcal{T}|}(s)$: value function of the final time.

Output: p_t^* : Optimal TCL outputs $p_{i,t}$ of each zone i at each time t .

- 1 **Offline reformulation:**
- 2 Conduct thermal dynamics partitioning to obtain independent models (14a) for each equivalent zone;
- 3 Convert chance constraints into deterministic constraints and transfer the optimization problem from (9) to (22);
- 4 For each equivalent zone c , work backward to quantify value functions $V_{i,t}$ from $t = |\mathcal{T}|$ to $t = 1$ recursively based on (24) in parallel;
- 5 **Online implementation:**
- 6 Initialize the state $s_{c,0}$ for each equivalent zone;
- 7 **for each period** $t \in \mathcal{T}$ **do**
- 8 Receive RTP λ_t from the electricity market;
- 9 **for each equivalent zone** $c \in \mathcal{C}$ **do**
- 10 Input $s_{c,t-1}^*$, λ_t , and $V_{c,t}(s_{c,t})$;
- 11 Solve (23) to obtain solutions $p_{c,t}, s_{c,t}$;
- 12 Recover control inputs of each zone $p_{i,t}^*, \forall i \in \mathcal{Z}$ based on $p_{c,t}, \forall c \in \mathcal{C}$;
- 13 Apply $p_{i,t}^*$ for TCL scheduling and measure the actual state of $s_{i,t}^*$ at time t ;
- 14 Collect $s_{i,t}^*$ and re-aggregate it to $s_{c,t}^*$ based on (14b) for scheduling.

coupled, and deterministic one, which is suitable for online TCL scheduling. The overall algorithm is presented in Algorithm 2. The offline stage includes sequentially thermal dynamics partitioning, converting chance constraints, and quantifying value functions. At each time stage in the online stage, we first make decisions for each equivalent single-zone case based on the actual real-time price, and then recover the decisions into the original building zones for TCL scheduling. Then, the real state can be obtained, and we re-aggregate the real states into equivalent states as the initial state for next-time scheduling.

Discussion: When applying the proposed methods in the practical scenarios, the proposed optimization framework is primarily formulated for variable-speed HVAC systems, where the heating/cooling power can be continuously modulated within the allowable range. Besides, the current work assumes ideal execution of the optimal decision on TCL power, meaning that the commanded value is accurately delivered at each time step without control delay or error. Furthermore, as the power output is optimized as the control input and the indoor temperature is the state variable, the current setting is closer to actuator control than conventional setpoint control.

V. CASE STUDIES

A. Experiment Setups

For algorithm implementation, we formulate a synthetic 100-zone office-building prototype for case studies. This prototype includes five floors with 20 zones per floor, as shown in Fig. 6. We then adopt a simplified RC-based thermal-dynamics model for TCL scheduling in this building. The RC values are the same for each floor, and the values at each floor are presented in Table II. The units of R and C are K/kW and kWh/K, respectively. The RC parameters are assumed to be known before the offline partitioning procedures. Their values are representative scenario parameters selected with reference to established equivalent thermal parameter (ETP) modeling studies [29], [30], in which equivalent resistances and heat capacities are derived from building geometry, envelope thermal properties, infiltration, and thermal mass. Here, R_i is the equivalent resistance of the aggregate exterior heat-transfer path, including envelope, glazing, infiltration, and ventilation effects, whereas R_{ij} is the equivalent resistance of the shared boundary and air-exchange path between zones i and j . Both are area-integrated equivalent quantities rather than material insulation values. In addition, C_i is an effective lumped capacitance including the active thermal mass of the building fabric and furnishings, rather than the heat capacity of the zone air alone.

The thermal disturbance data are provided by the “Building Energy Codes Program” in the U.S. Department of Energy [31]. They comprise δ_t from the outdoor temperature and $\delta'_{i,t}$ from the combined heat gains of occupants, lighting, and electric equipment. Specifically, outdoor temperature data are collected from Typical Meteorological Year (TMY) files. Occupancy-related gains are generated from the EnergyPlus objects, scaling with the effective occupant number. The number is calculated as the design number times the scheduled occupancy fraction, corresponding to approximately 59.4–237.5 W heat gains per zone when occupied. Heat gains from solar radiation are not included. This is a modeling-scope simplification that is consistent with reduced thermal models in related studies [32], [33], rather than an assumption that solar heat gain is physically negligible. We present both the outdoor temperature data and internal disturbance (e.g., occupancy heat gains) from sampled 10 days in February in Fig. 5.

Besides, we collect RTP and DAP data from the New York Independent System Operator (NYISO) as input for optimization. The time interval for demand response is set as 5 minutes, which is the same as RTP. This chosen time interval is appropriate for both the execution time of control commands and the variation of indoor temperature. On the one hand, it is assumed that there is an ideal execution of the optimized power command without actuator delay. On the other hand, different from set-point control, we take the HVAC power as the control input and regard the indoor temperature as the state, which does not require indoor temperatures to reach a new steady state within a 5-minute interval. Other parameters for datasets, zone partitioning, and optimization are concluded in Table III. Specifically, the value of $\rho_{i,t}$ remains unchanged across different zones and time slots. It is reasonable because

Table II
RC VALUES OF THE BUILDING THERMAL DYNAMICS MODEL

Zone (i)	R_i	C_i	Zone (ij)	R_{ij}	Zone (ij)	R_{ij}
1	10.08	10.50	1,2	15.00	10,12	13.50
2	6.48	9.60	2,3	12.00	11,12	2.24
3	10.08	9.90	2,11	1.28	11,14	12.00
4	5.40	7.50	2,14	14.00	11,15	12.00
5	5.40	7.50	3,4	12.50	11,18	3.52
6	/	7.20	3,6	12.50	11,19	3.52
7	/	6.90	3,11	12.50	12,13	15.00
8	10.08	11.40	4,5	2.88	12,19	3.68
9	7.20	7.20	4,6	2.88	12,20	12.00
10	7.20	7.20	5,7	2.88	14,15	3.52
11	/	12.60	5,8	12.50	14,16	3.52
12	6.48	9.00	6,7	2.88	15,16	13.50
13	10.08	10.20	6,11	12.50	15,17	13.50
14	10.08	13.20	7,8	12.50	15,18	13.50
15	/	9.60	7,11	12.50	16,17	2.88
16	5.40	7.50	8,9	12.50	17,20	2.88
17	5.40	7.50	8,11	14.00	18,19	2.56
18	/	7.20	9,10	2.72	18,20	3.04
19	/	7.20	9,11	13.00	19,20	3.04
20	7.92	7.80	10,11	13.00	/	/

Table III
PARAMETER SETTINGS

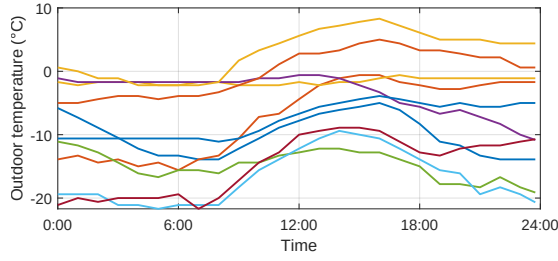
Parameters	Configuration
Dataset	Electricity price: NYISO
	Building prototype: 100-zone office
	Period: 2023/01-2023/02
Partitioning	Cluster number ($ \mathcal{C} $): 8
	Minimum size ($\underline{M}$): 1
	Maximum size ($\overline{M}$): 7
Optimization	Time interval: 5 min
	Risk level (γ): 0.05
	Power limit (P): [0, 4kW]
	Model error: $\mu = 0, \sigma = 0.01$
	($\underline{S}, \overline{S}$) for zones 1,6-8,13-17: [21°C, 25°C]
	($\underline{S}, \overline{S}$) for zones 2,11-12,20: [20°C, 26°C]
	($\underline{S}, \overline{S}$) for zones 3-5,9-10,18-19: [20.5°C, 25.5°C]
	Comfort penalty coefficient ρ : 0.008
	Price forecasting sample number: 20
	Value function sample number: 800

the zone-level and temporal heterogeneity are mainly represented by comfort bounds. Furthermore, the value of $\rho_{i,t}$ is set as 0.008, and the value is not a physical parameter, but a manually-set weighting parameter in the quadratic comfort penalty function. The basic principle of valuation is to balance the relative magnitude between the economic cost of electricity consumption and the penalty for temperature deviation. An ideal value is to ensure that comfort violations are sufficiently penalized while not completely dominating the price-driven TCL scheduling. All algorithms are executed in MATLAB on a computer with a 2.60 GHz Intel 13th CPU with 16 GB of RAM. Source code, input data, and solving results are available on Github¹. More information regarding the synthetic building benchmark and other case study details can be found in [34].

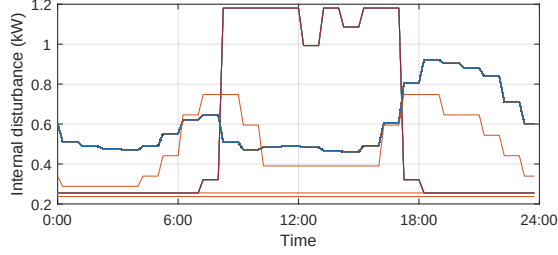
B. Optimization Results

1) *Comparisons*: To verify the effectiveness of this proposed method in reducing daily scheduling costs and mitigating comfort violations, this subsection presents several

¹<https://github.com/hkuedl/Online-Multi-zone-TCL-Scheduling>



(a) External disturbance (outdoor temperature)



(b) Internal disturbance

Figure 5. Disturbance values across sampled 10 days in February.

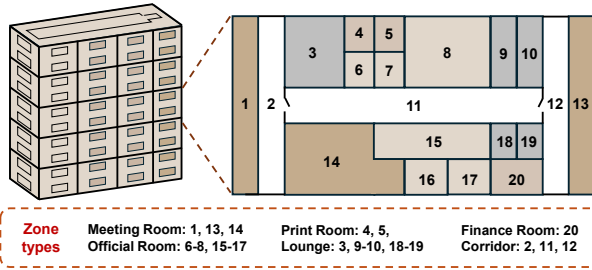


Figure 6. Building prototype.

methods for performance comparison. The introduction of all comparison methods is concluded as follows:

- **Ground-Truth (GT):** This method integrates the ground-truth model with perfect price forecasts to solve the deterministic optimization problem, and the solved results of this ideal case serve as globally-optimal solutions for reference.
- **DEcoupled-Analytical Stochastic Dynamic Programming (DE-ASDP):** Compared with the proposed method, the difference of this method is that it employs a decoupled model for optimization, where thermal interactions between these 100 zones are fully ignored. Additionally, this method employs the same two-stage deterministic transformation to solve the problem.
- **PRoposed-Stochastic Model Predictive Control (PR-SMPC):** Compared with the proposed method, the difference of this method is that it employs stochastic MPC (SMPC) for handling the decision deviations caused by both model and price forecast errors, while it utilizes the same thermal dynamics partitioning method to derive the required model. Key parameters include a 2-h prediction horizon for MPC with 5-min time step (i.e., a total of 24 stages), and a 5-sample uncertainty set at each step for

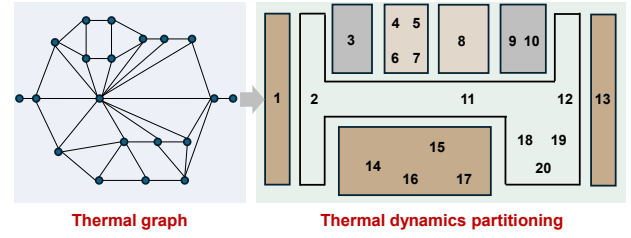


Figure 7. Thermal dynamics partitioning results.

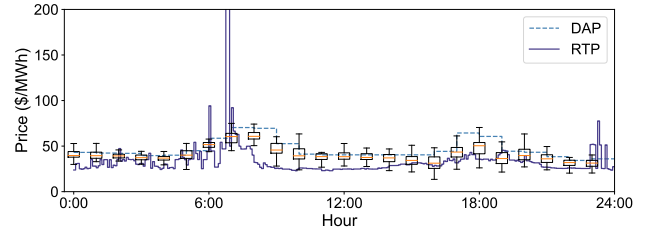


Figure 8. Real-time and day-ahead prices.

stochastic optimization.

- **PRoposed-Multi-resolution Model Predictive Control (PR-MMPC):** Compared with PR-SMPC, the difference of this method is that it employs multi-horizon MPC (MMPC), while it utilizes the same thermal dynamics partitioning method. In MMPC, finer time steps are used for near-term scheduling decisions, whereas coarser time steps are used for distant prediction periods to reduce decision stages. Key parameters include a 2 h prediction horizon with 5-min time steps for the first 20 min and 50-min time steps for the latter 120 min (i.e., a total of 6 stages), and a 5-sample uncertainty set for stochastic optimization.
- **PRoposed-Analytical Stochastic Dynamic Programming (PR-ASDP) (Proposed):** The proposed method with parameter settings in Subsection V.A.

2) *Optimization costs:* We first perform clustering-based thermal dynamics partitioning for the 100-zone building used. We set the cluster number as 8 for each floor (40 zones in total), and the partitioning results are given in Fig. 7. Given the results, it can be seen that neighboring zones with similar structures and strong thermal interactions are classified into the same cluster, which can then be further aggregated into 8 equivalent single zones without coupled interactions. Besides, we set error distributions between RTP and DAP in 2023/01 and impose that on the revealed DAP in 2023/02 as RTP forecasts. The difference between real-time and day-ahead prices (DAP) is shown in Fig. 8, where RTP exhibits more volatile, frequent fluctuations than DAP, leaving more room for price arbitrage in TCL scheduling.

Given the above settings, we implement online TCL scheduling during 10 sampled days within 2023/02. The daily optimization costs, including the economic costs of power purchase and the comfort penalty, after online scheduling, are shown in Fig. 9. The average and the standard deviation values of the 10-day optimization results are summarized in

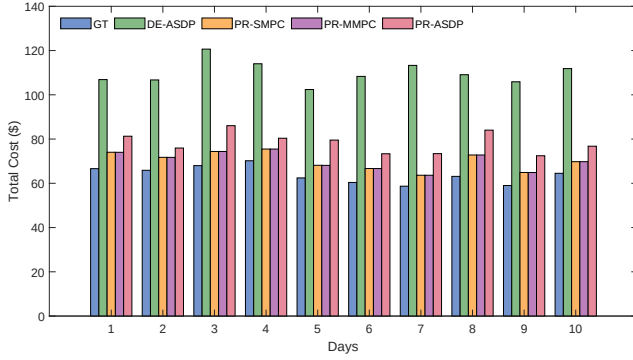


Figure 9. Optimization costs of TCL scheduling in 10 days.

Table IV
COST DISTRIBUTION OF TCL SCHEDULING OVER 10 DAYS

	GT	DE-ASDP	PR-SMPC	PR-MMPC	PR-ASDP
Total cost	63.89 ± 3.86	109.91 ± 5.20	70.16 ± 4.18	70.15 ± 4.18	78.32 ± 4.69
Power cost	19.11 ± 3.99	19.64 ± 3.98	20.35 ± 4.22	20.35 ± 4.22	18.65 ± 3.97
Comfort penalty	44.77 ± 0.40	90.27 ± 4.87	49.81 ± 0.31	49.80 ± 0.31	59.67 ± 3.89

Table IV. It can be found that PR-SMPC, PR-MMPC, and PR-ASDP obtain similar optimization costs that are close to the ground-truth results. Specifically, the proposed PR-ASDP method is less economically efficient with a relatively larger comfort penalty than the PR-SMPC method, while PR-SMPC and PR-MMPC have quite similar cost results. Thus, the proposed deterministic optimization method achieves comparable performance to SMPC in PR-SMPC, demonstrating its effectiveness in obtaining accurate decisions in deterministic optimization. In contrast, the DE-ASDP method, with a fully decoupled model as input, generally achieves the highest scheduling cost. The high cost of DE-ASDP is mainly due to the comfort penalty term, indicating that the decoupled model ignores all thermal interactions and results in much higher model errors than other methods.

Specifically, the comfort penalty is the major item that causes performance gaps among different methods. We select the fourth day as a specific case and compare the comfort penalties on one floor (i.e., 20 zones) in Fig. 10. DE-ASDP therefore produces the largest comfort penalty, particularly in zones 6, 7, 15, 18, and 19. By retaining the principal inter-zone thermal interactions, the three partitioning-based methods remain closer to GT. In particular, PR-ASDP reduces the comfort penalty by approximately 35.9% relative to DE-ASDP.

3) *Computational time*: Besides, the operational efficiency is also important for real-time decision-making in online TCL scheduling. We present the computational time of all comparison methods on each day in Table V. It can be seen that the proposed ASDP method in both DE-ASDP and PR-ASDP can make decisions in less than 1s, while the computational time of PR-SMPC ranges from 0.4h to 3h. The major advantage of ASDP is that it only needs to solve the single-stage, deterministic problem that most existing commercial solvers can handle. On this basis, the decoupled model used in these methods enables parallel computation

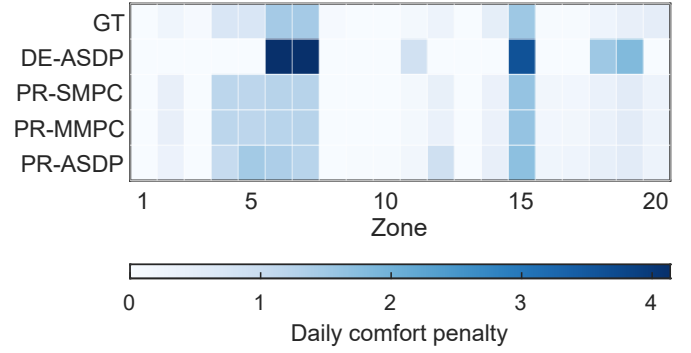


Figure 10. Comfort penalties of all methods on the fourth day of study.

Table V
COMPARISON OF COMPUTATIONAL TIME (s)

Days	DE-ASDP	PR-SMPC	PR-MMPC	PR-ASDP
1	1.17	1220.14	983.07	0.23
2	0.85	2685.72	2455.00	0.23
3	0.60	4139.48	3924.49	0.23
4	0.82	5632.60	5395.21	0.23
5	0.94	7291.76	6906.76	0.24
6	0.66	9161.37	8779.56	0.23
7	0.53	880.25	1023.74	0.23
8	0.57	2143.88	2543.83	0.23
9	0.52	4325.88	4047.39	0.23
10	0.75	5805.60	5535.01	0.23

within each zone, allowing individual decisions. Note that the proposed PR-ASDP has more stable computational time because it solves the problem with fewer zones (40) than the original DE-ASDP model (100). By contrast, at each time stage, the traditional SMPC must solve a multi-stage, sample-based stochastic problem over the forecasting horizon using multiple samples. The computational time makes online scheduling less feasible than ASDP. A similar problem can be found in PR-MMPC. In MMPC, less computational time is required than in SMPC by reducing the total prediction stages, but such a decrease is marginal because it still requires solving the sample-based stochastic optimization problem. Besides, the computational time of both SMPC and MMPC is unstable across days and highly dependent on the forecasting samples of the electricity price. Note that the scheduling-cost and computational-time performance are mainly based on the relative relationship across inter-zone thermal interactions, showing less sensitivity to the relative magnitude between R_i and R_{ij} .

C. Algorithm Performance

In this subsection, we further verify the algorithm's performance with respect to two technical details in the proposed method: the sensitivity analysis of the clustering number and the priority of ASDP for real-time optimization.

1) *Sensitivity analysis of clustering*: We take 2023/02/01 as the optimization case, and we solve the daily optimization with perfect forecasting of the electricity price. We set different models for comparison: utilizing the actual model to represent the ground-truth results, the fully decoupled model to represent the decoupled results, and the thermal dynamics partitioning

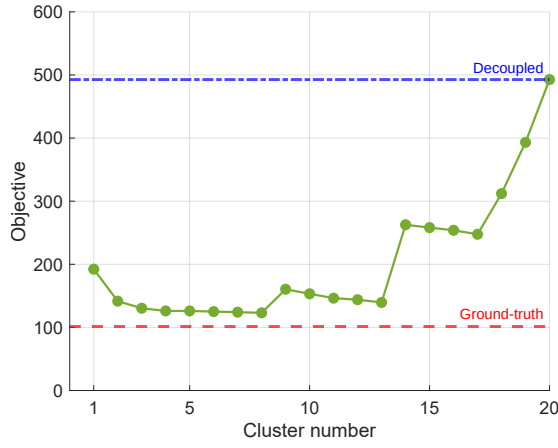


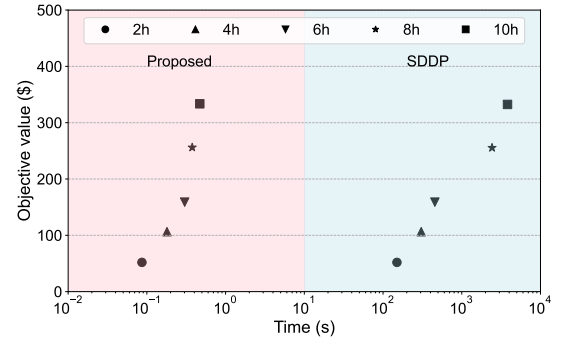
Figure 11. Sensitivity analysis on clustering numbers.

to represent the results under different clustering numbers. The sensitivity analysis on cluster numbers is presented in Fig. 11. The x-axis represents the number of equivalent zones in each floor. It can be seen that the methods reach relatively low optimization costs with 3-13 clusters. The cost increases with more clusters because more thermal interactions among neighboring zones are wrongly mitigated. The results are equivalent to decoupled ones when the cluster number equals 0. The cost also increases with fewer clusters because the thermal interactions are wrongly clustered and the aggregation errors increase.

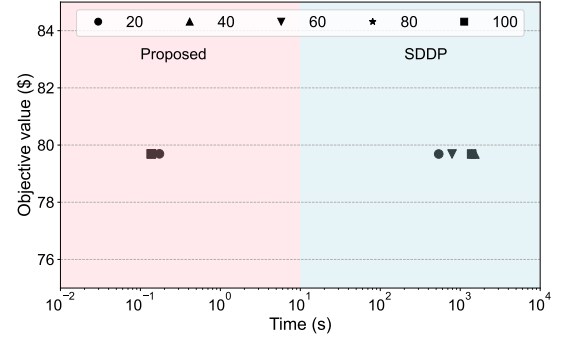
2) *Effectiveness of ASDP*: To verify the performance of ASDP in terms of computational efficiency and solution accuracy, we compare the proposed ASDP with stochastic dual dynamic programming (SDDP), implemented in the FAST toolbox [35]. The SDDP solver simulates 100 Monte-Carlo scenarios at each forward pass with 20 maximum iterations. For the proposed method, the value function is sampled with 500 state samples from -10°C - 10°C at each time stage. We model the RTP forecast as zero-mean normal distributions over DAP on 2023/02/01, with a standard deviation of 20.

We first set different time lengths ranging from 2h-10h (24-120 stages) with 30 samples in the price distribution. For example, there will be a total of 30^{24} scenarios for 2h optimization. The comparisons of computational time and objective values are presented in Fig. 12a. The results show that the computational time of SDDP increases significantly with the time length, whereas the proposed method finishes within several seconds, up to 10^3 to 10^4 times faster than SDDP. Note that both methods achieve similar objective values, showing that the proposed method can obtain value functions as accurate as SDDP in a very short time. Besides, we also set different price samples ranging from 20-100 with the same time length (3h). The results in Fig. 12b show that the computational efficiency of the proposed method is almost agnostic to the size of samples, achieving real-time response within 1s across all cases.

3) *Robustness analysis to disturbances*: Here we verify the robustness of the proposed work to different disturbance cases. Specifically, we consider the fluctuation of ambient temperature δ and the extreme weather cases within the



(a) Different stages



(b) Different samples

Figure 12. Comparison of computational efficiency.

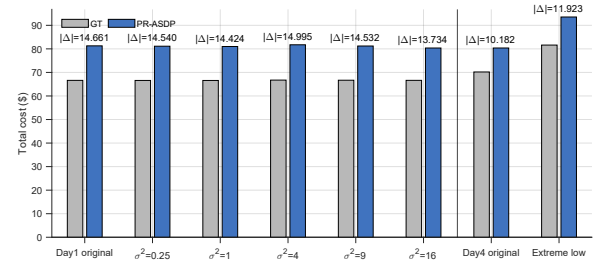


Figure 13. Accuracy comparison of objective costs under various disturbances.

disturbance term in the model. For the fluctuation of ambient temperature, we update the original disturbance term as $\delta_t^{\text{up}} = \delta_t + \epsilon_t$, $\epsilon_t \in \mathcal{N}(0, \sigma^2)$, where ϵ_t is the added disturbance following the normal distribution $\mathcal{N}$ with different variance values σ . We set five cases of $\sigma = [0.5, 1.0, 2.0, 3.0, 4.0]$ to represent different fluctuating degrees. Besides, we set day 4 as an extreme low-temperature case by reducing the temperature by 10°C at all time steps. The comprehensive comparison results are presented in Fig. 13. It can be found that the absolute errors between PR-ASDP (the proposed method) and GT remain similar under different σ and extremely low weather. Therefore, the performance of the proposed work is robust against different patterns of ambient weather.

4) *Robustness analysis to measurement errors*: Here we further verify the robustness of the proposed work to different measurement errors. In (4), we have already considered the stochastic model errors. Here, random noise is introduced into the measured indoor temperature (i.e., Lines 13-14 at

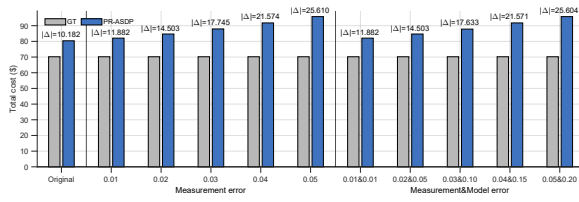


Figure 14. Accuracy comparison of objective costs under measurement and model errors.

Algorithm 2) during online scheduling. The proposed method is then evaluated under different noise levels and compared with the ground-truth solution. The noise levels are set as $[0.10, 0.20, 0.30, 0.40, 0.50]$. The first five studies separately test the influence of measurement noises, and the following five studies test the coupled influence of both measurement noises and model errors (σ in (4b) is set as $[0.01, 0.05, 0.10, 0.15, 0.20]$). The results are presented in Fig. 14. It can be seen that the performance of the proposed method is more sensitive to the measurement error than to the model error. When the noise level rises from 0.01 to 0.05, the absolute error of the objective cost increases from 11.882 to 25.610. At each time step, the method could make a deviated evaluation of the initial value because of measurement errors, and such time-coupled deviations could amplify the decision errors in the following steps, resulting in relatively larger errors. In contrast, the model error has been fully considered as chance constraints (16), which could have a slight influence on decision errors when the chance constraint is not active after solving.

D. Scalability Analysis

In this subsection, we evaluate scalability under different numbers of zones and heterogeneous thermal characteristics. To avoid changing the problem size by simply duplicating thermally independent floors, the number of floors is fixed at five in all cases, while the number of zones on each floor is increased from 20 to 40. For the 40-zone case, a second vertically mirrored 20-zone prototype is joined to the first along a common boundary through ten explicit inter-zone thermal resistances, connecting zones 1, 2, 3, 4, 5, 8, 9, 10, 12, and 13 to their mirrored counterparts. The nominal interface resistance is 7.2 K/kW. Thus, each 40-zone floor forms one connected thermal network rather than two independent copies. The four cases are defined as follows:

- SC-100 (Strong Coupling–100 zones): Five floors with 20 zones per floor. The inter-zone thermal resistances are scaled by 0.8 with random perturbations.
- SC-200 (Strong Coupling–200 zones): Five floors with 40 thermally connected zones per floor. The inter-zone thermal resistances are scaled by 0.8 with random perturbations.
- WC-100 (Weak Coupling–100 zones): Five floors with 20 zones per floor. The inter-zone thermal resistances are scaled by 1.2 with random perturbations.
- WC-200 (Weak Coupling–200 zones): Five floors with 40 thermally connected zones per floor. The inter-zone thermal resistances are scaled by 1.2 with random perturbations.

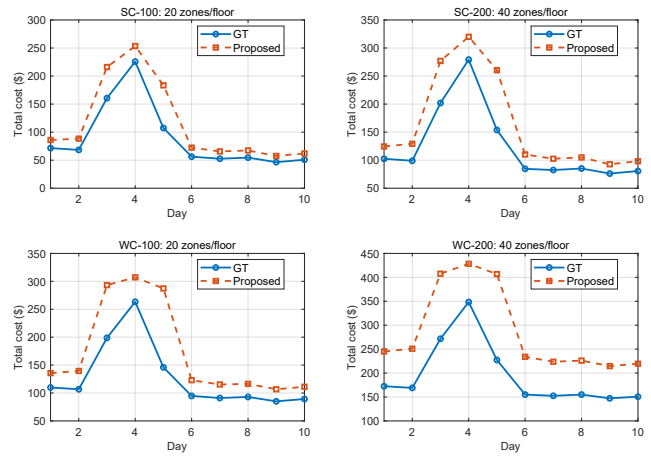


Figure 15. Cost comparison of TCL scheduling under different buildings.

The ten-day cost comparisons between the ground truth (GT) and the proposed method are presented in Fig. 15. For SC-100, SC-200, WC-100, and WC-200, the mean GT total costs (\$) are 89.45, 124.56, 127.76, and 194.93, respectively, while the mean relative cost gaps are 29.0%, 29.4%, 34.7%, and 47.7%, respectively. When the number of thermally connected zones is doubled, the cost gap therefore remains almost unchanged under strong coupling. The larger cost gap observed for WC-200 indicates that, under weak thermal coupling, the partitioned approximation becomes more sensitive as the connected thermal network grows.

VI. CONCLUSIONS

This work follows a stepwise approach to reformulate the original coupled, stochastic, and multi-stage optimization problem into a decoupled, deterministic, and single-stage one sequentially, enabling the problem to be applicable for online TCL scheduling. The proposed methods leverage a clustering-based partitioning method to aggregate the zones with high thermal interactions. It is helpful to maintain thermal interactions across specific zones while decoupling specific zones to obtain equivalent single and independent zones for facilitating parallel computing. Compared with the decoupled model without assuming thermal interactions, the proposed method could avoid around 33.9% of the comfort penalty caused by model errors. For each decoupled zone, the following two-stage deterministic transformation method converts the stochastic and chance-constrained problem into one with a deterministic objective function and constraints, making it compelling for solving with existing solvers. Compared with the traditional solving of stochastic problems, it helps save around 10^4 - 10^5 computational time for efficient solving. Meanwhile, the loss of scheduling cost is less than 11.5%.

In future research, we will further investigate the performance of the proposed method under heterogeneous building structures and extended time periods by incorporating the seasonal pattern of ambient temperature into the model. Besides, more realistic scenarios, including on/off control pattern, actual control error or delay, and diverse control logic, will be considered in future applications.

APPENDIX A

PROOF OF PROPOSITION 2

We apply dual decomposition with dual variable θ_t associated with (14a), and the problem (23) is decomposed as

$$\min_{p_t} \{\lambda_t p_t - \beta \theta_t p_t \mid \text{s.t. (22c)}\}, \quad (25a)$$

$$\min_{s_t} \{U_t(s_t) + V_t(s_t) + \theta_t s_t \mid \text{s.t. (16)}\}. \quad (25b)$$

The dual updating with step size $\epsilon \in \mathbb{R}^+$ is:

$$(1/\epsilon)\dot{\theta}_{t-1} = -s_t + \eta s_{t-1} + \beta p_t + \varepsilon_t \quad (26)$$

The solution to (25) satisfies

$$p_t = \begin{cases} P & \text{if } \lambda_t < \beta \theta_t \\ 0 & \text{else} \end{cases}, \quad (27a)$$

$$s_t = \begin{cases} w_t^{-1}(-\theta_t) & \text{if } -w_t(\bar{S}_t - \bar{\xi}_t) \leq \theta_t \leq -w_t(\underline{S}_t - \underline{\xi}_t) \\ \bar{S}_t - \bar{\xi}_t & \text{if } \theta_t < -w_t(\bar{S}_t - \bar{\xi}_t) \\ \underline{S}_t - \underline{\xi}_t & \text{if } \theta_t > -w_t(\underline{S}_t - \underline{\xi}_t) \end{cases} \quad (27b)$$

where we denote $w_t = v_t + u_t$. Note that w is monotonic increasing as u, v are both monotonic increasing. Thus, the s has a unique solution regarding θ .

The following aim is to derive $v_{t-1}(s_{t-1})$, which can be written as

$$\begin{aligned} v_{t-1}(s_{t-1}) &= \mathbb{E}[q_{t-1}(s_{t-1})] = \int_{-\infty}^{\infty} x \cdot \Pr[q_{t-1} = x] dx \\ &= \int_{-\infty}^{\infty} x \cdot \frac{\partial \Pr[q_{t-1}(s_{t-1}) \leq x]}{\partial x} dx. \end{aligned} \quad (28)$$

Regarding θ_t^* as optimal solutions, we have to derive

$$\Pr[q_{t-1}(s_{t-1}) \leq x] = \Pr[-\eta \cdot \theta_t^* \leq x] = \Pr[\theta_t^* \geq -\frac{x}{\eta}], \quad (29)$$

We denote $y = -\frac{x}{\eta}$. For $y \leq \theta_t^*$, we have the following relationship based on the iterative condition of θ :

$$s_t(y) \leq \eta s_{t-1} + \beta p_t(y) + \varepsilon_t, \quad (30)$$

where $s_t(y), p_t(y)$ are solutions to (25) when the dual variable is set to y .

We denote the right hand of (30) as $s_t^P(y)$, which can be re-written by (27a) as

$$s_t^P(y) = \begin{cases} \eta s_{t-1} + \beta P + \varepsilon_t' & \text{if } \lambda_t < \beta y \\ \eta s_{t-1} + \varepsilon_t' & \text{else} \end{cases} \quad (31)$$

Then we have

$$\begin{aligned} \Pr[s_t(y) \leq s_t^P(y)] &= \Pr[w_t(s_t(y)) \leq w_t(s_t^P(y))] \\ &= \Pr[y \geq -w_t(s_t^P(y))] \end{aligned} \quad (32)$$

It can be seen that

$$\begin{aligned} \Pr[q_{t-1}(s_{t-1}) \leq x] &= \Pr[y \geq -w_t(s_t^P(y))] \\ &= \Pr[\lambda_t < \beta y] \cdot \mathbf{1}[y \geq -w_t(\eta s_{t-1} + \beta P + \varepsilon_t')] \\ &\quad + \Pr[\lambda_t \geq \beta y] \cdot \mathbf{1}[y \geq -w_t(\eta s_{t-1} + \varepsilon_t')] \end{aligned} \quad (33)$$

After ignoring the subscription, (33) can be represented as

$$\Pr[q_{t-1}(s) \leq x] = \begin{cases} 0 & \text{if } y < -w_t(\eta s + \beta P + \varepsilon') \\ F(\beta y) & \text{if } -w_t(\eta s + \beta P + \varepsilon') \leq y < -w_t(\eta s + \varepsilon') \\ 1 & \text{if } y \geq -w_t(\eta s + \varepsilon') \end{cases} \quad (34)$$

Correspondingly, we have

$$\Pr[q_{t-1}(s) = x] = \begin{cases} F(-\frac{\beta x}{\eta}) & \text{if } x = \eta \cdot w_t(\eta s + \beta P + \varepsilon') \\ -\frac{\beta}{\eta} f(-\frac{\beta x}{\eta}) & \text{if } \eta w_t(\eta s + \varepsilon') < x < \eta w_t(\eta s + \beta P + \varepsilon') \\ 1 - F(-\frac{\beta x}{\eta}) & \text{if } x = \eta \cdot w_t(\eta s + \varepsilon') \\ 0 & \text{else} \end{cases} \quad (35)$$

where the discontinuous points $x = \eta w_t(\eta s + \varepsilon')$ and $x = \eta w_t(\eta s + \beta P + \varepsilon')$, their probabilities are the difference between the two neighboring cumulative distribution results:

$$\Pr[q_{t-1}(s) = x] = \Pr[q_{t-1}(s) \leq x] - \Pr[q_{t-1}(s) < x] \quad (36)$$

Substituting (35) back into (28), we complete the proof to derive $v_{t-1}(s_{t-1})$ in (24).

REFERENCES

- [1] IEA, "Heating," [Online], <https://www.iea.org/energy-system/buildings/heating>.
- [2] X. Liu, X. Liu, Y. Jiang, T. Zhang, and B. Hao, "Photovoltaics and energy storage integrated flexible direct current distribution systems of buildings: Definition, technology review, and application," *CSEE Journal of Power and Energy Systems*, vol. 9, no. 3, pp. 829–845, 2023.
- [3] California Energy Commission, "Thermal zones, HVAC zones and space functions," [Online], <https://energycodeace.com/site/custom/public/reference-ace-2013/index.html#!Documents/22thermalzoneshvaczonesandspacefunctions.htm>.
- [4] B. He, N. Zhang, C. Fang, Y. Su, and Y. Wang, "Flexible building energy management with neural odes-based model predictive control," *IEEE Transactions on Smart Grid*, vol. 15, no. 5, pp. 4690–4704, 2024.
- [5] Q. Chen and N. Li, "Fast simulation and high-fidelity reduced-order model of the multi-zone radiant floor system for efficient application to model predictive control," *Energy and Buildings*, vol. 248, p. 111210, 2021.
- [6] X. Cui, Y. Wang, and B. Xu, "Dimension-reduced optimization of multi-zone thermostatically controlled loads," *IEEE Transactions on Smart Grid*, vol. 16, no. 6, pp. 4685–4697, 2025.
- [7] J. Joe, P. Im, B. Cui, and J. Dong, "Model-based predictive control of multi-zone commercial building with a lumped building modelling approach," *Energy*, vol. 263, p. 125494, 2023.
- [8] Q. Lin, L. Ding, Z. Kong, Z.-W. Yu, X. Li, and H. Wang, "Multi-time scale model predictive control-based demand side management for a microgrid," *IEEE Transactions on Smart Grid*, vol. 16, no. 2, pp. 1181–1193, 2025.
- [9] L. Yu, Y. Sun, Z. Xu, C. Shen, D. Yue, T. Jiang, and X. Guan, "Multi-agent deep reinforcement learning for HVAC control in commercial buildings," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 407–419, 2021.
- [10] C. P. Barala, A. Firdous, P. Mathuria, M. Shahidehpour, and R. Bhakar, "Virtual energy storage system framework for multi-temperature zone buildings," *IEEE Transactions on Industry Applications*, 2026.
- [11] Q. Zhu, Y. Yang, L. Yu, Q.-S. Jia, C. J. Spanos, and X. Guan, "An equivalent and unified virtual battery modeling framework for flexibility characterization of building HVAC systems," *arXiv preprint arXiv:2512.21363*, 2025.
- [12] D. H. Green, Y. Lin, A. Botterud, J. Gregory, S. B. Leeb, and L. K. Norford, "Pareto-optimized thermal control of multi-zone buildings using limited sensor measurements," *IEEE Transactions on Smart Grid*, vol. 15, no. 5, pp. 4674–4689, 2024.

- [13] Y. Zhou, Y. Jiang, and H. Zhao, "A fitting-based modeling method of TCLs for power system demand response," in *2025 10th International Conference on Power and Renewable Energy (ICPRE)*. IEEE, 2025, pp. 81–84.
- [14] Y.-J. Kim, "A supervised-learning-based strategy for optimal demand response of an HVAC system in a multi-zone office building," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4212–4226, 2020.
- [15] V. Taboga and H. Dagdougui, "A distributed ADMM-based deep learning approach for thermal control in multi-zone buildings under demand response events," *IEEE Transactions on Automation Science and Engineering*, 2024.
- [16] W. Li, S. Wang, and C. Koo, "A real-time optimal control strategy for multi-zone VAV air-conditioning systems adopting a multi-agent based distributed optimization method," *Applied Energy*, vol. 287, p. 116605, 2021.
- [17] M. Mork, A. Xhonneux, and D. Müller, "Nonlinear distributed model predictive control for multi-zone building energy systems," *Energy and Buildings*, vol. 264, p. 112066, 2022.
- [18] Y. Yang, J. Shi, D. Wang, C. Wu, and Z. Han, "Carbon-aware scheduling of thermostatically controlled loads: A bilevel DRCC approach," *IEEE Transactions on Smart Grid*, vol. 16, no. 2, pp. 1233–1247, 2025.
- [19] N. S. Raman, R. U. Chaturvedi, Z. Guo, and P. Barooah, "Model predictive control-based hierarchical control of a multi-zone commercial HVAC system," *Journal of Engineering for Sustainable Buildings and Cities*, vol. 2, no. 2, p. 021005, 2021.
- [20] G. Chen, H. Zhang, H. Hui, and Y. Song, "Fast wasserstein-distance-based distributionally robust chance-constrained power dispatch for multi-zone HVAC systems," *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 4016–4028, 2021.
- [21] W. Xue, N. Jia, and M. Zhao, "Multi-agent deep reinforcement learning based HVAC control for multi-zone buildings considering zone-energy-allocation optimization," *Energy and Buildings*, vol. 329, p. 115241, 2025.
- [22] Y. Xia, X. Wang, X. Yin, W. Bo, L. Wang, S. Li, and K. Li, "Federated accelerated deep reinforcement learning for multi-zone HVAC control in commercial buildings," *IEEE Transactions on Smart Grid*, 2025.
- [23] R. Liu, Y. Pan, and Y. Chen, "Laxity-aware scalable reinforcement learning for HVAC control," *arXiv preprint arXiv:2306.16619*, 2023.
- [24] P. Mohebi, S. Li, and Z. Wang, "Chance-constrained stochastic framework for building thermal control under forecast uncertainties," *Energy and Buildings*, vol. 331, p. 115385, 2025.
- [25] E. Atam and E. C. Kerrigan, "Optimal partitioning of multithermal zone buildings for decentralized control," *IEEE Transactions on Control of Network Systems*, vol. 8, no. 3, pp. 1540–1551, 2021.
- [26] J. Drgoňa, M. Kvasnica, M. Klaučo, and M. Fikar, "Explicit stochastic mpc approach to building temperature control," in *52nd IEEE conference on decision and control*. IEEE, 2013, pp. 6440–6445.
- [27] L. Chen and B. Xu, "A prudent framework for understanding risk-awareness in demand response," *European Journal of Operational Research*, 2025.
- [28] B. Xu, M. Korpås, and A. Botterud, "Operational valuation of energy storage under multi-stage price uncertainties," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 55–60.
- [29] Z. Zheng, S. Wang, W. Li, and X. Luo, "A consensus-based distributed temperature priority control of air conditioner clusters for voltage regulation in distribution networks," *IEEE Transactions on Smart Grid*, vol. 14, no. 1, pp. 290–301, 2022.
- [30] C. D. Goodman, L. E. Hinkle, T. D. Hardy, and H. M. Reeve, "Building model calibration: validation of GridLAB-D thermal dynamics modeling," Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), Tech. Rep., 2022.
- [31] The U.S. Department of Energy (DOE), "Prototype building models," [Online], <https://www.energycodes.gov/prototype-building-models>.
- [32] X. Chen, Y. Li, J. Shimada, and N. Li, "Online learning and distributed control for residential demand response," *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 4843–4853, 2021.
- [33] R. Liu, H. Hui, X. Chen, and Y. Song, "Distributed frequency control of heterogeneous energy storage systems considering short-term ability and long-term flexibility," *IEEE Transactions on Smart Grid*, vol. 15, no. 6, pp. 5693–5705, 2024.
- [34] "Supplementary information of this study," [Online], <https://drive.google.com/file/d/12gzA2YsPFEFBgoJ3QIM3LWY95yycGRWk/view?usp=sharing>.
- [35] "Fast (finally an sddp toolbox)," [Online], <https://stanford.edu/~lcmambier/cgi-bin/fast/index.php>.



Xueyuan Cui received the Ph.D. degree from The University of Hong Kong, Hong Kong SAR, China, in August 2025. He received the B.E. and M.S. degrees from Zhejiang University, Hangzhou, China, in June 2019 and March 2022, respectively. He is currently a Postdoctoral Fellow in the Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University. His research interests include data-driven building energy modeling and optimization.



clusiveness.

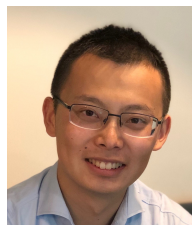
Liudong Chen is a Ph.D. candidate in Earth and Environmental Engineering at Columbia University. He was a Ph.D. summer intern at the National Laboratory of the Rockies (NLR) in 2023 and at ISO New England in 2025. He was the recipient of the IEEE Power & Energy Society (PES) Outstanding Student Scholarship in 2022. His research focuses on analyzing and modeling behavior in energy systems through optimization, economics, and machine learning, addressing the challenges related to increasing uncertainty, interdependence, and in-



include data analytics in smart grids, energy forecasting, multi-energy systems, Internet-of-things, and cyber-physical-social energy systems.

Yi Wang received the B.S. degree from Huazhong University of Science and Technology in June 2014, and the Ph.D. degree from Tsinghua University in January 2019. He was a visiting student with the University of Washington from March 2017 to April 2018. He served as a Postdoctoral Researcher in the Power Systems Laboratory, ETH Zurich from February 2019 to August 2021.

He is currently an Assistant Professor with the Department of Electrical and Electronic Engineering, The University of Hong Kong. His research interests



energy storage, power system optimization, and power system economics.

Bolun Xu received the B.S. degree in electrical engineering from Shanghai Jiaotong University, Shanghai, China, in 2011, the M.Sc. degree in electrical engineering in the Swiss Federal Institute of Technology, Zurich, Switzerland, in 2014, and the Ph.D. degree in electrical engineering from the University of Washington, Seattle, WA, USA, in 2018. He is currently an Assistant Professor of Electrical Engineering with the Department of Earth and Environmental Engineering, Columbia University, New York, NY, USA. His research interests include