An RC-Network Approach for HVAC Precooling Optimization in Buildings

Hongsen Shi, Jia Liu, *Senior Member, IEEE,* and Qian Chen

Abstract—To lower buildings' significant energy consumption and high impacts on environmental sustainability, recent years have witnessed rapidly growing interests in efficient HVAC precooling control and optimization. However, due to the complex analytical modeling of building thermal transfer, rigorous mathematical optimization for HVAC precooling is highly challenging. As a result, progress on HVAC precooling optimization remains limited in the literature. Our main contribution is that we overcome the aforementioned challenge and propose an accurate and tractable HVAC precooling optimization framework. The main results of this paper are three-fold: i) We develop an RC-network-based analytical model for multi-zone HVAC precooling to minimize both total energy costs and peak load demand; ii) We show that the HVAC procooling optimization problem based on the proposed RC network model admits a convex approximation, which enables an efficient optimization algorithm design; and iii) Based on the convex approximation insight and by exploiting special problem structures, we develop an efficient distributed algorithm to solve the HVAC precooling optimization problem. Further, we conduct extensive simulation studies to verify the performance of our proposed mathematical model and algorithms. Our numerical results indicate that compared with the five existing HVAC control strategies, the proposed algorithm consistently outperforms existing state-of-the-art approaches.

Index Terms—Energy efficiency, building HVAC precool scheduling, distributed optimization, algorithms

1 Introduction

Buildings have a significant impact on the global climate change and other energy-related environmental issues [1]. In the United States, over one-third of all energy and around 70% of electricity were consumed by buildings [2]. Further, according to U.S. Department of Energy (USDOE), around 40% of the total energy used by buildings is consumed by the heating, ventilation, and air conditioning (HVAC) systems [3]. In particular, during hot summer months, cooling requirements would skyrocket and become the dominant source of HVAC energy consumption. Exacerbating the problem is the fact that these cooling requirements correlate strongly both in space and time (e.g., a period with peak solar radiation at a certain geographical region, the time-span during which most commercial buildings are fully occupied, etc.), which causes a dramatic surge in the peak power consumption in the grid. Such a sharply increased energy demand necessitates ramping up uneconomical and pollutive generators, further creating serious environmental sustainability concerns. Therefore, to avoid exceedingly high peak loads, many electric utility companies have imposed heavy price penalties during the peak load period. For most commercial buildings, the electricity bills typically contain two parts: total energy consumption charges and peak demand charges (i.e., calculated based on the maximum energy demand in kW in a certain peak demand interval, e.g., 15 minutes, during the billing cycle [4]). Although pricing strategies vary with service providers, peak demand charges always occupy a significant proportion of the electricity bills, sometimes even exceeding 50% [5]. As a result, during the peak time when buildings are fully occupied and operational, there is an inherent dilemma between procuring energy at extremely high costs and maintaining indoor thermal comfort, an important metric for indoor environmental quality.

Clearly, to resolve the peak-time dilemma, one has to shift the HVAC energy demands away from the peak time period in some way. One effective approach for shifting the HVAC peak load is to leverage building thermal mass properties to perform *HVAC precooling*. Simply speaking, the basic idea behind precooling is that, if a building has been properly precooled before occupancy or during the early morning occupied time, then even if the HVAC is turned off, the building temperature will not jump immediately and would gradually and slowly rise up thanks to the "memory effect" of the building thermal mass. As a result, an acceptable low temperature would maintain and last into the peak time period, which helps reduce the peak time HVAC energy demand. Moreover, precooling during off-peak times is often assisted by the relatively low ambient temperature (e.g., late nights or early mornings) and cheaper time-of-use electricity rates, which would unlikely incur a high energy demand or a higher energy consumption charge. Because of this win-win situation, there has been a great deal of interest in developing optimal HVAC precooling control strategies to lower buildings' peak demand as well as reduce the total electricity bill.

However, performing optimal HVAC precooling is

- *Manuscript received December 15, 2018; revised August 22, 2019; accepted September 17, 2019. J. Liu's work is supported in part by NSF Grants CNS-1758757, CCF-1758736, ECCS-1818791, ONR N00014-17-1-2417, and AFRL FA8750-18-1-0107. Q. Chen's work is supported in part by NSF Grant CNS-1446582.*
- *H. Shi and Q. Chen are with the Department of Food, Agriculture, and Biological Engineering, The Ohio State University, Columbus, OH 43210.E-mail: {shi.546, chen.1399}@osu.edu*
- *J. Liu is with the Department of Computer Science, Iowa State University, Ames, IA 50011.E-mail: jialiu@iastate.edu*

highly non-trivial. The main reason is that there is a lack of an *accurate* and *tractable* theoretical framework to enable rigorous mathematical optimization for HVAC precooling. To date, most building energy simulation tools based on the DOE-2 engine [6] (e.g., EnergyPlus, eQuest, etc.) calculate through-wall heat transfer by solving a complex second-order ordinary differential equation (ODE). This ODE takes into account a large number of practical building factors, such as local weather, building geometry, building envelope characteristics, internal heat gains from lighting, people and plug loads, HVAC system specifications, etc. Although being detailed and sophisticated, the high-complexity of the DOE-2 model renders it hopeless to derive closed-form analytical expressions to formulate tractable optimization problems. In fact, even for the simpler resistive-capacitative (RC) thermal transfer model (based on a simplified first-order ODE, see Section 3 for details), it remains too complex to be used as a starting point for designing optimization algorithms. Due to these challenges, many existing studies in HVAC precooling scheduling (see, e.g., [7], [8], [9], [10], [11]) are limited to either heuristics or simple single-zone settings (see Section 2 for more detailed discussions).

The major contribution of this paper is that we overcome the aforementioned challenges and develop an accurate and tractable mathematical model for rigorous HVAC precooling optimization. The main technical results in this paper are as follows:

- Based on an RC-based network model for multi-zone HVAC, we develop an accurate and tractable multi-objective mathematical optimization framework for HVAC precooling that considers both total energy costs and peak load demand. Specifically, by exploiting the finite dynamic range characteristic of HVAC systems, we show that the thermal energy transfer model based on the RC-network model can be closely approximated and linearized. As a result, the original HVAC precooling optimization problem based on the RC thermal dynamics ODE can be converted into a *convex approximation*, which enables efficient optimization design. We further show that the error of this convex approximation can be made arbitrarily small as the number of time-slots in the system grows asymptotically, hence offering a graceful trade-off between energy cost optimality and the problem dimension complexity.
- Based on our developed convex approximation above, we exploit its separable problem structure to propose an efficient ADMM-type (alternating direction method of multipliers) distributed algorithm for solving the HVAC precooling problem. Moreover, we show that the special structural properties of the augmented Lagrangian naturally imply low-complexity and efficient computational schemes for the primal temperature setpoints and HVAC energy injection decisions. Specifically, we prove that the temperature setpoint in each time-slot can be computed in a "backward induction" fashion. Also, by recognizing an interesting *rank-1* correction structure, we show that the HVAC control decision variables in each time-slot can be efficiently computed by leveraging the Sherman-Morrison-Woodbury (SMW) matrix inversion technique. We note that these insights are not only elegant mathematically, but they also lead to highly efficient HVAC control protocol designs in practice.
- To verify the performance of our proposed mathematical model and algorithms, we conduct extensive and in-depth simulation studies based on a large number of floor plans and internal structures. The effects of five building parameters (i.e., the gross floor area, average room size, total number of zooms, wall capacitance, and the window-to-wall ratio), internal loads, and external environments on the optimal strategy are investigated based on daily simulations. We also compare our optimization algorithm with five existing strategies (i.e., the occupancy-driven ON/OFF strategy and four other strategies combining precooling with different heuristic demand-limiting control methods) in terms of cooling energy cost reduction. Overall, we show that our algorithm consistently performs better by achieving cooling energy cost reduction ratios ranging up to 75% when using these existing strategies as the baseline, thus confirming the efficacy of our proposed mathematical optimization framework and algorithmic design.

Collectively, our results in this paper contribute to a comprehensive and fundamental understanding of the roles of HVAC precooling optimization on environmental sustainability. The remainder of the paper is organized as follows. Section 2 reviews the related work about the existing precooling HVAC control strategies and the RC network. Section 3 introduces the RC network model and the problem formulation. Section 4 focuses on problem reformulation and linearized approximation, which further motivates efficient algorithm designs in Section 5. Section 6 presents simulations and numerical studies and Section 7 concludes this paper.

2 Related Work

In the literature, it has long been recognized that a building's thermal mass holds a great potential to shift the building's HVAC loads and reduce the peak demand [9], [12], [13], [14]. However, as mentioned in Section 1, due to the lack of an accurate and tractable theoretical optimization framework, progress on HVAC precooling optimization has been quite limited. As a result, most existing studies on HVAC precooling strategies design resort to simple heuristics. For example, three HVAC precool scheduling schemes termed concave-increasing, step-up, and linear-up were used in [8], [9] to gradually reduce the HVAC energy input during the peak time. Another more sophisticated heuristic approach is based on the model-based predictive control (MPC) (also known as receding horizon control) [15], [16], [17]. Simply speaking, MPC is a method of process control that uses empirical dynamic models combined with future predictions to perform optimization in the current time-slot [10]. However, it is usually unclear how to determine a proper mathematical model for HVAC, which is a key part of the MPC approach [11]. Although shown to be effective in varying degrees, the major limitation of these heuristic approaches is that there is a lack of optimality guarantee.

To put HVAC precooling optimization on a firmer analytical footing, in recent years, there have been several lines

of research on developing mathematical optimization techniques for HVAC precooling. For example, Lee and Braun in [18] proposed three optimization methods to determine the trajectories of setpoint temperature during the peak time, including the semi-analytical (SA) method, the exponential setpoint equation-based semi-analytical (ESA) method, and the load weighted-average (WA) method. In their follow-up work [19], these three methods were further combined with precooling or extended precooling to reduce the peak cooling load. While these three methods can achieve optimality guarantee during the peak time to some extent, they do not consider cooling schedules during the off-peak time. This is in stark contrast to our work, where we *jointly* optimize cooling schedules for both peak and non-peak periods.

The most related line of work to our paper is based on the first-order RC model [20], [21], [22]. Simply speaking, an RC model captures the physical properties of walls in a building to predict thermal transfer transients, which is widely popular in computing cooling or heating requirements for performance monitoring, diagnosis, and control strategy analysis [23] (see Section 3 for further background of RC). Based on the RC model, Mukherjee et. al. [20] developed a thermal feedback control scheme for a multi-zone building, and Bhattacharya et. al. [22] proposed a dynamic programming (DP) based algorithm to solve a nonlinear optimal precooling control problem in buildings under the time-of-use electricity pricing. However, the DP solution is only limited to the single-zone setting and cannot be extended to multi-zone HVAC due to the curse of dimensionality of DP. For RC-based multi-zone HVAC optimization, Gupta et al. [21] proposed a two-stage approach, where a consensus-based algorithm is first used to solve a static optimization problem to obtain optimal stationary states for all time periods. Then, they designed a control law in each time period to drive the system to reach the desired stationary states. However, due to the high-complexity, this two-stage approach is cumbersome to implement in practice. By contrast, in this paper, we propose a new convex approximation for RC thermal transfer ODE to avoid the pitfalls of [21], [22]: On one hand, the low-complexity of our approach allows us to handle multi-zone settings; on the other hand, the linearized approximation directly captures the transient dynamics in the systems, thus eliminating the need for the two-stage process in [21] and enabling efficient algorithmic designs.

3 System Model and Problem Formulation

In this paper, we use boldface to denote matrices/vectors. We let $\mathbf{A}^\top$ denote the transpose of $\mathbf{A}$. We let $(\mathbf{A})_{ij}$ represent the entry in the i-th row and j-th column of $\mathbf{A}$ and let $(\mathbf{v})_m$ represent the m-th entry of $\mathbf{v}$. We let $\mathbf{I}$ and $\mathbf{O}$ denote the identity and all-zero matrices, respectively, where their dimensions are conformal to the context. We let $\mathbf{1}$ and $\mathbf{0}$ denote the all-one and all-zero vectors, respectively, where their dimensions are conformal to the context.

1) HVAC System Modeling: Consider a building with N HVAC zones that are indexed by $i \in \{1, \ldots, N\}$. For building safety concerns and to maintain an appropriate human comfort zone, the temperature of each zone i must be maintained within a range $[T_i^{lb}, T_i^{ub}]$ at all times, where T_i^{lb} and T_i^{ub} represent the lower and upper temperature bounds of zone i, respectively. We consider summer operations, where the ambient temperature is typically higher than the indoor temperature. As a result, air conditioning (AC) operation suffices and heating is not needed. Our goals in this paper are two-fold: i) Understanding whether precooling this N-zone building under time-varying electricity prices can save cost; and ii) If yes, how to design an optimal precool schedule to minimize the cooling energy cost.

To this end, we assume a look-ahead time window $[0, W]$ (e.g., a day), for which accurate electricity price and weather forecasts are available. The look-ahead time window is equally divided into K time-slots, which are indexed by $k \in \{1, \ldots, K\}$. Hence, the duration of each time-slot can be computed as $\tau = W/K$. We assume K to be large enough (or equivalently, τ is sufficiently short) such that the electricity price and the ambient temperature in each time-slot remain static. Thus, we let $p[k]$ and $T_A[k]$ denote the electricity price, and ambient temperature in time-slot k, respectively,

We let $T_i[k]$ denote the indoor temperature setpoint of zone i at the beginning of the k-th time-slot. Clearly, the collection of all setpoints $\{T_i[k], \forall i, k\}$ constitutes a *cooling schedule*. Correspondingly, we let $u_i[k]$ represent the cooling energy consumption in zone i and time-slot k that achieves the temperature setpoints $T_i[k]$[1]. As will be seen later, our precooling schedule optimization amounts to determining the values of $\{T_i[k], u_i[k], \forall i, k\}$.

2) Resistive-Capacitative (RC) Thermal Dynamics Model: In this paper, we consider a heat transfer model based on thermal resistance and capacitance, which has been widely adopted in civil and mechanical research communities. Previous research has evaluated the accuracy of RC models of varied complexity (e.g., 2R1C, 3R2C, etc.) and found that RC model-based simulation can achieve reasonable accuracy in temperature prediction with the root-mean-square error of around 0.5 °C [24], [25] and building energy consumption prediction with the relative error within 10% [26]. In what follows, we first give a primer on the fundamentals of the RC model to familiarize the readers with the necessary background.

In the RC model for a given building, each zone (e.g., a room, a hall way, etc.) is modeled as a thermal capacitor and each wall is modeled as a concatenation of $n+1$ thermal resistors and n thermal capacitors, $n \geq 1$. Simply speaking, thermal resistance models the thermal energy flow based on temperature difference: $Q = \Delta T/R$, where Q is the thermal energy (in unit W) transferred across the resistance, ΔT is the temperature difference (in unit K), and R is the thermal resistance (in unit K/W). On the other hand, thermal capacitance models the ability of space/mass to store heat: $C\frac{dT}{dt} = Q$, where C has the unit J/K. In practice, the most widely used RC model is the 3R2C model, i.e., $n = 2$. As shown in Figure 1, under the 3R2C model, the wall separating two HVAC zones i and j in a building is

1. In practice, the AC power output levels are usually discrete. Assume that the AC system for each zone i has a cooling power rating of $\bar{u}_i$. Then, any AC power output value $0 \leq u_i[k] \leq \bar{u}_i$ can be attained by duty cycling the HVAC compressor between ON (power output $\bar{u}_i$) and OFF (power output 0) states with low amplitude and high frequency [22].

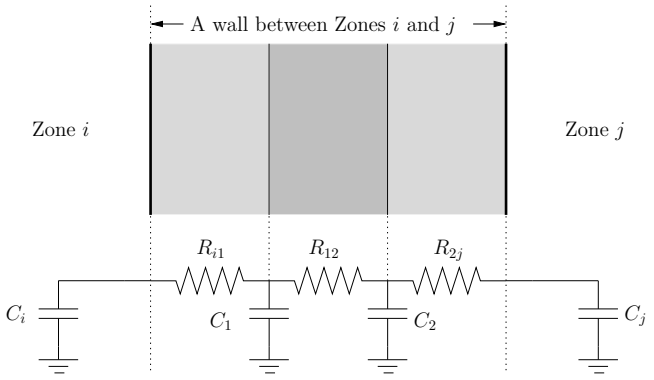


Fig. 1. The 3R2C model for heat transfer across a wall.

composed of three thermal resistors (R_{i1}, R_{12}, R_{2j}) and two thermal capacitors (C_1, C_2). Also, zone i and zone j are modeled as two thermal capacitors C_i and C_j, respectively.

It can be seen that the RC model is analogous to an electric circuit. As a result, the thermal dynamic under the RC model is also closely related to the classical circuit theory. More specifically, let $S_i^{[k]}(t)$ denote the temperature of the i-th thermal capacitor at time instant t within time-slot k. We note that it is important to distinguish the two notions of "time" introduced so far. Earlier in this paper, we have used bracket "$[k]$" to denote a time-slot k (in a larger time-scale); while in here, we use parenthesis "(t)" to signify a time instant within some given time-slot. Under the RC model, the evolution of $S_i^{[k]}(t)$ is governed by the following first-order ordinary differential equation (ODE):

$$C_i \frac{dS_i^{[k]}(t)}{dt} = \sum_{j \in \mathcal{N}_i} \frac{S_j^{[k]}(t) - S_i^{[k]}(t)}{R_{ij}} + Q_{Ai}^{[k]}(t) + Q_{Hi}^{[k]}(t) + Q_{Li}^{[k]}(t), \tag{1}$$

where $\mathcal{N}_i$ denotes the set of thermal capacitors connected to thermal capacitor i, R_{ij} denotes the thermal resistor between zone i and zone j (we assume $R_{ij} = R_{ji}$), and $Q_{A,i}^{[k]}(t)$ is the instantaneous thermal power transferred from the ambient environment to thermal capacitor i at time instant t in time-slot k and defined as follows:

$$Q_{Ai}^{[k]}(t) = \begin{cases} \frac{T_A[k] - S_i^{[k]}(t)}{R_{0i}}, & \text{if capacitor } i \text{ is adjacent to the ambient,} \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

where R_{0i} represents the thermal resistor between capacitor i and the ambient environment. Note that $Q_{Ai}^{[k]}(t)$ can implicitly capture the combined effect of some green building technologies. For example, cool air in early morning can be injected into the building as free cooling input in some green buildings. In (1), $Q_{Hi}^{[k]}(t)$ denotes the cooling power input injected from the HVAC system and defined as follows:

$$Q_{Hi}^{[k]}(t) = \begin{cases} u_i^{[k]}(t), & \text{if capacitor } i \text{ is connected with HVAC,} \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

where $u_i^{[k]}(t)$ denotes the instantaneous HVAC cooling power control decision to zone i at time t in time-slot k. In (1), $Q_{Li}^{[k]}(t)$ denotes the internal loads (mainly including plug loads, lighting load, and occupancy load) in zone i and time-slot k, which are defined as follows:

$$Q_{Li}^{[k]}(t) = \begin{cases} o_i^{[k]}(t), & \text{if capacitor } i \text{ has internal loads,} \\ 0, & \text{otherwise,} \end{cases} \tag{4}$$

where $o_i^{[k]}(t)$ denotes the instantaneous internal loads in zone i at time t in time-slot k. We note that the prediction of internal load $Q_{Li}^{[k]}(t)$ can be accurately made based on the building's schedules and/or usage statistics in the past that depend on the day of week, month, or season, etc. For example, the occupancy of academic buildings can be easily predicted based on class schedules, the occupancy of office buildings can be predicted between weekdays and weekends, the occupancy of a convention center can be predicted based on its events schedule, etc.

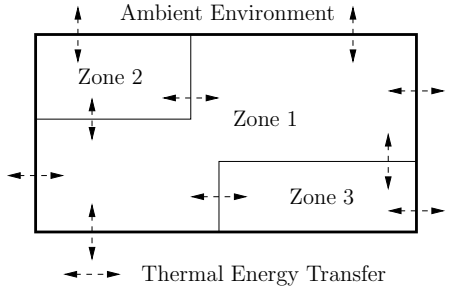


Fig. 2. The layout of a three-zone building example.

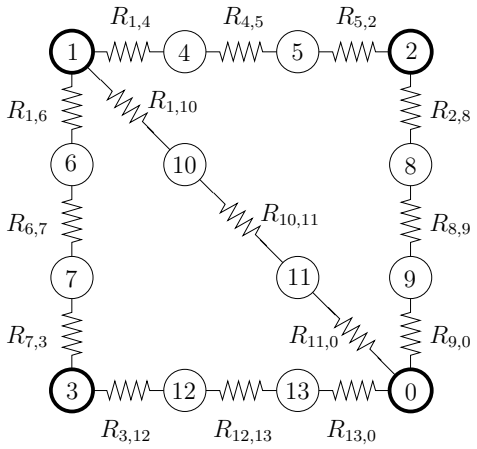


Fig. 3. The RC network for the three-zone building example in Figure 2.

3) RC Network: With the above RC-based thermal dynamics modeling, we are now in a position to use an RC network to model the thermal transfer in a building with a multi-zone HVAC system. In this paper, a building is viewed as a connected network $\mathcal{G} = \{\mathcal{N}, \mathcal{L}\}$, where $\mathcal{N}$ and $\mathcal{L}$ denote the set of nodes and links, respectively. Each node in $\mathcal{N}$ corresponds to a thermal capacitor. Also, a reference node is added to represent the ambient environment. We note that all thermal capacitors are modeled as nodes in the network, including both HVAC zones and wall capacitors in the RC wall model. Each link in $\mathcal{L}$ represents a thermal resistor with two end nodes corresponding to the two adjacent thermal capacitors. We let N and L denote the total numbers of thermal capacitors and resistors in the

network, respectively. Hence, the network has $N+1$ nodes (including the ambient environment as the reference node) and L links. For convenience, we label the nodes from 0 to N, with node 0 denoting the ambient environment. R_{ij} denotes the thermal resistor between capacitors i and j.

As an example, Figures 2 and 3 illustrate the layout of a three-zone building example and its corresponding 3R2C network. In the RC network, the thermal capacitors 1, 2, 3, and 0 in Figure 3 (denoted by bolded circles) correspond to zones 1, 2, 3, and the ambient environment in Figure 2, respectively. Thermal capacitors 4–13 represent the capacitors in the 3R2C models for walls. In Figure 2, it can be seen that zones 1 and 2 are connected in the sense that they are adjacent and separated by a wall. Thus, in Figure 3, a 3R2C model connection is used to represent the wall between zone 1 and zone 2: three thermal resistors $R_{1,6}$, $R_{6,7}$, $R_{7,3}$, and two thermal capacitors C_6 and C_7. Other wall resistors and capacitors in Figure 3 can also be identified following the same token.

4) Problem Formulation: In this paper, our goal is to optimize the cooling schedule to minimize the total electricity energy expense combined with the peak load demand. As noted in the introduction, the strong correlation of cooling requirements in both space and time causes a dramatic surge in the peak power consumption in the grid, which necessitates ramping up uneconomical and pollutive generators. Therefore, we focus on jointly optimizing energy cost and peak demands in this paper. Let Γ denote the set of time-slots in the peak period. Based on the modeling described earlier, the objective function can be computed as $\sum_{k=1}^{K}\sum_{i=1}^{N}\int_{t=0}^{\tau} p[k]\mathbf{u}_i^{[k]}(t)dt + \hat{P}(\max_{k\in\Gamma}\int_0^{\tau} u_i^{[k]}(t)dt)$, where $\hat{P}(\cdot)$ denotes the peak load penalty function. In this paper, we assume that $\hat{P}(\cdot)$ is an increasing strongly convex function. Putting together all analytical modeling above, we can write the HVAC precooling optimization (HPrO) problem as follows:

HPrO:

$$\underset{u_i^{[k]}(t),\forall i,k,t}{\text{Minimize}} \quad w_1\sum_{k=1}^{K}\sum_{i=1}^{N}\int_{t=0}^{\tau} p[k]u_i^{[k]}(t)dt + w_2\hat{P}\left(\max_{k\in\Gamma}\sum_{i=1}^{N}\int_0^{\tau} u_i^{[k]}(t)dt\right) \tag{5}$$

$$\text{subject to} \quad C_i\frac{dS_i^{[k]}(t)}{dt} = \sum_{j\in\mathcal{N}_i}\frac{S_j^{[k]}(t)-S_i^{[k]}(t)}{R_{ij}} + \frac{T_A[k]-S_i^{[k]}(t)}{C_iR_{Ai}}\mathbb{1}_A(i) + u_i^{[k]}(t)\mathbb{1}_H(i) + o_i^{[k]}(t)\mathbb{1}_L(i),$$
$$i=1,\ldots,N, t\in[0,\tau], k=1\ldots,K, \tag{6}$$

$$S_i^{[k]}(0) = T_i[k],\ i=1,\ldots,N,\ k=1,\ldots,K, \tag{7}$$

$$T_i^{lb} \le S_i^{[k]}(t) \le T_i^{ub}, \quad \forall i,k,t, \tag{8}$$

$$u_i^{[k]}(t) \in \mathcal{U}_{i,t}^{[k]}, \quad \forall i,k,t, \tag{9}$$

where $\mathcal{U}_{i,t}^{[k]}$ represents the set of all feasible HVAC control decisions for $u_i(t)$ at time t; $\mathbb{1}_A(i)$ is an indicator function that takes value 1 if zone i is connected to the ambient environment and 0 otherwise; $\mathbb{1}_H(i)$ is an indicator function that takes value 1 if zone i is directly connected to the HVAC system and 0 otherwise; and $\mathbb{1}_L(i)$ is an indicator function that takes value 1 if zone i has internal loads and 0 otherwise. In (5), the weights w_1 and w_2 are non-negative constants with $w_1+w_2=1$, which represents the relative emphasis on minimizing total energy cost and peak load demand, respectively. The extreme cases $(w_1=1, w_2=0)$ and $(w_1=0, w_2=1)$ correspond to minimizing total energy cost and minimizing peak load demand, respectively. In Problem HPrO, Eq. (6) represents the RC-based indoor temperature evolution dynamics with (7) being the initial condition, and Eq. (8) represents the human comfort zone constraints.

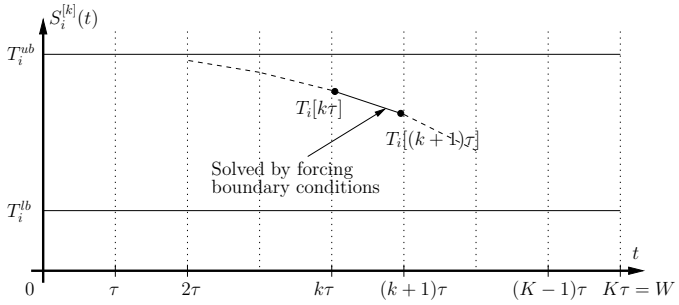


Fig. 4. A snapshot of the approximatly linear behavior of temperature $S_i^{[k]}(t)$ as the duration τ gets small.

We note that Problem HPrO is a continuous-time optimization problem that is difficult to solve directly in a computationally tractable fashion using conventional optimization techniques. This is mainly due to the existence of the integration of an *unknown* function $u_i^{[k]}(t)$ in the objective function (5) and in the ODE of the temperature dynamics in (6). As a result, conventional nonlinear optimization techniques cannot be directly applied. In the next section, we will show how these challenges can be addressed by a linear reformulation of Problem HPrO.

4 Reformulation for the HVAC Precooling Optimization Problem

In this section, we will propose a convex approximation approach to reformulate Problem HPrO, which leads to a convex programming problem and hence a tractable optimization solution. Moreover, we will show that the error of the proposed convex approximation approach can be made arbitrarily small by trading off complexity in problem size. The basic idea behind our approximation and reformulation approach is to let the duration of a time-slot tend to be infinitesimal (i.e., letting $\tau\to 0$), so that the objective function in (5) and (6) can be linearized and convexified. In what follows, we will demonstrate the key steps and components of our convex approximation and reformulation approach.

Step 1): Convexifying the HVAC Precooling Objective: As mentioned earlier, one of the major hurdles in solving Problem HPrO is the integration of the HVAC input decisions, which cannot be handled by standard optimization techniques. What is worse is that the objective function of Problem HPrO involves an unknown decision function $u_i^{[k]}(t)$, which further makes the problem intractable. Our key to approximate and convexify the objective function

in (5) is to exploit the *physical characteristics* of most HVAC systems that the cooling power input can only be changed gradually. More specifically, the HVAC system control usually does not allow sudden jumps to avoid inefficiency or damages to its electronics components. Therefore, as the time-slot duration τ gets small, the function $u_i^{[k]}(t)$ can be viewed as a constant between the time-slot interval $[0, \tau]$. Let $u_i[k]$ denote the constant energy input from the HVAC in the k-th time-slot for zone i. Hence, the objective function of Problem HPrO in (5) can be reformulated as:

$$(5) \approx \underset{u_i[k],\forall i,k}{\text{Min}} \; w_1 \sum_{k=1}^{K}\sum_{i=1}^{N} p[k]u_i[k] + w_2 \hat{P}\left(\max_{k\in\Gamma} \sum_{i=1}^{N} u_i[k]\right), \quad (10)$$

if τ is sufficiently small. Also, since we have assumed that the ambient temperature and the electricity price can be viewed as constants (cf. Section 3) if the time-slot duration τ is sufficiently small, further reducing τ will not violate these assumptions.

Step 2): Convexifying the RC thermal dynamics ODE: Next, we turn our attention to the RC thermal dynamics ODE in (6). Note that due to the capacitive thermal mass in the building, the temperature curve $S_i^{[k]}(t)$ in each zone i also exhibits gradual changes. Hence, as τ gets small, the $S_i^{[k]}(t)$ function can be well approximated by a linear function that passes through the setpoints $T_i[k]$ and $T_i[k+1]$, as illustrated in Figure 4. As the time-slot duration τ gets small, the function $o_i^{[k]}(t)$ can also be viewed as a constant between the time-slot interval $[0, \tau]$. Let $o_i[k]$ denote the constant internal loads in the k-th time-slot for zone i. Based on these observations, we can develop a linear approximation for the RC thermal dynamics ODE as follows. First, we let $\mathcal{N}_{i0} \triangleq \mathcal{N} \cup \{0\}$ (i.e., the set combining node i's neighbors and the ambient environment) and define $a_i \triangleq \sum_{j\in\mathcal{N}_{i0}} (C_i R_{ij})^{-1}$, $i = 1, \ldots, N$, where we let $R_{i0} \triangleq \infty$, $\forall i$. Next, we define the following constants:

$$p_i \triangleq e^{a_i\tau}, \qquad i \in \{1, \ldots, N\}, \quad (11)$$

$$p_{ij} \triangleq \frac{(a_i\tau - 1)e^{a_i\tau} + 1}{a_i^2 C_i R_{ij}\tau}, \; i, j \in \{1, \ldots, N\}, \quad (12)$$

$$q_{ij} \triangleq \frac{a_i\tau - e^{a_i\tau} + 1}{a_i^2 C_i R_{ij}\tau}, \qquad i, j \in \{1, \ldots, N\}, \quad (13)$$

$$r_i \triangleq \frac{e^{a_i\tau} - 1}{a_i\tau}, \qquad i \in \{1, \ldots, N\}, \quad (14)$$

$$s_i \triangleq \frac{e^{a_i\tau} - 1}{a_i C_i R_{0i}}, \qquad i \in \{1, \ldots, N\}. \quad (15)$$

Then, we can show the following important ODE linearization result:

Theorem 1 (RC Thermal Transfer ODE Linearization)**.** *Let the time-slot duration τ be sufficiently small such that: i) $u_i^{[k]}(t)$ and $o_i^{[k]}(t)$ are approximately static with $u_i^{[k]}(t) = u_i[k]/\tau$ and $o_i^{[k]}(t) = o_i[k]/\tau, \forall i, k, t$; and ii) the temperature evolution curve of each zone i in all time-slots k is approximately a line segment connecting $T_i[k]$ and $T_i[k+1]$. Then, the RC thermal dynamics ODE in (6) can be linearized as:*

$$p_i T_i[k+1] - \sum_{j\in\mathcal{N}_i} p_{ij} T_j[k+1] - T_i[k] + \sum_{j\in\mathcal{N}_i} q_{ij} T_j[k] - r_i u_i[k]\mathbb{1}_H(i) - r_i o_i[k]\mathbb{1}_L(i) = s_i T_A[k], \quad i = 1, \ldots, N, k = 1, 2, \ldots, K-1. \quad (16)$$

Theorem 1 can be proved by solving the ODE under the stated assumptions and we relegate the details to Appendix A.

Step 3): A Convex Programming Reformulation: Based on the previous two steps, it can be seen that both the objective function and the RC thermal dynamics ODE have been approximated by a sum of a convex function and a linear function concerning the HVAC inputs $u_i[k]$ and temperature setpoints $T_i[k]$, $\forall i, k$ (i.e., precooling schedules), respectively. Hence, we can approximate and reformulate the original Problem HPrO as a convex optimization problem as follows:

R-HPrO:

$$\underset{u_i[k],\forall i,k}{\text{Minimize}} \; w_1 \sum_{k=1}^{K}\sum_{i=1}^{N} p[k]u_i[k] + w_2 \hat{P}(z) \quad (17)$$

$$\text{subject to } z \geq \sum_{i=1}^{N} u_i[k], \; \forall k \in \Gamma, \quad (18)$$

$$p_i T_i[k+1] - \sum_{j\in\mathcal{N}_i} p_{ij} T_j[k+1] - T_i[k] + \sum_{j\in\mathcal{N}_i} q_{ij} T_j[k] - r_i u_i[k]\mathbb{1}_H(i) - r_i o_i[k]\mathbb{1}_L(i) = s_i T_A[k], \quad i = 1, \ldots, N, k = 1, 2, \ldots, K-1, \quad (19)$$

$$T_i^{lb} \leq T_i[k] \leq T_i^{ub}, \; i = 1, \ldots, N, k = 1, \ldots, K, \quad (20)$$

$$u_i[k] \in [0, u^{\max}], \; i = 1, \ldots, N, k = 1, \ldots, K, \quad (21)$$

where we introduce an auxiliary variable z to reformulate and simplify the minmax objective function. Two remarks regarding Problem R-HPrO are in order: First, it is clear that as $\tau \to 0$ (by letting the number of time-slots K go to infinity), the per-slot static HVAC input and linear temperature evolution approximations (cf. Theorem 1) can be made arbitrarily accurate. This means that the solution to the proposed reformulated Problem R-HPrO approaches to that of the original problem asymptotically by trading off complexity of the problem size (reflected in the number of time-slots). Also, it can be seen that Problem R-HPrO is a convex optimization problem since the objective function is convex and all constraints are linear. Thus, Problem R-HPrO can be solved in polynomial time by general interior-point method. However, we note that Problem R-HPrO possesses several interesting special structural properties, which can be exploited to enable even more efficient algorithm design, and more importantly, distributed control implementation. This constitutes the major subject in the next section.

5 Efficient Algorithm Design

Thanks to the RC network model, the reformulated Problem HPrO possesses a special network structure, which allows us to develop efficient algorithms. To see this, we first

introduce several matrix notations to further restate Problem R-HPrO in a more compact form as follows.

We start with restating the linearized RC thermal dynamics in (19). To this end, we let $\mathbf{T}[k] \triangleq [T_1[k] \ \dots \ T_N[k]]^\top \in \mathbb{R}^N$ be the vector that collects all temperature setpoints at the beginning of time-slot k. We also let $\mathbf{u}[k] \triangleq [u_1[k] \ \dots \ u_N[k]]^\top \in \mathbb{R}^N$ and $\mathbf{o}[k] \triangleq [o_1[k] \ \dots \ o_N[k]]^\top \in \mathbb{R}^N$ be the vectors that collect all HVAC energy inputs and the internal loads in time-slot k, respectively. Further, we let $\mathbf{T} \triangleq [\mathbf{T}[1] \ \dots \ \mathbf{T}[K]]^\top \in \mathbb{R}^{NK}$, $\mathbf{u} = [\mathbf{u}[1] \ \dots \ \mathbf{u}[K]]^\top \in \mathbb{R}^{NK}$, $\mathbf{o} = [\mathbf{o}[1] \ \dots \ \mathbf{o}[K]]^\top \in \mathbb{R}^{NK}$, $\mathbf{s} \triangleq [s_1 \ \dots \ s_N]^\top \in \mathbb{R}^N$, and $\mathbf{r} \triangleq [r_1 \mathbb{1}_L(1) \ \dots \ r_N \mathbb{1}_L(N)] \in \mathbb{R}^N$. Also, we define three matrices $\mathbf{D}$, $\mathbf{E}$, and $\mathbf{G}$ as follows:

$$[\mathbf{D}]_{ij} \triangleq \begin{cases} p_i = e^{a_i \tau}, & \text{if } j = i, \\ p_{ij} = -\frac{(a_i\tau - 1)e^{a_i\tau}+1}{a_i^2 C_i R_{ij} \tau}, & \text{if } j \neq i \text{ and } j \in \mathcal{N}_i, \\ 0, & \text{otherwise.} \end{cases} \tag{22}$$

$$[\mathbf{E}]_{ij} \triangleq \begin{cases} 1, & \text{if } j = i, \\ q_{ij} = -\frac{a_i\tau - e^{a_i\tau}}{a_i^2 C_i R_{ij} \tau}, & \text{if } j \neq i \text{ and } j \in \mathcal{N}_i, \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

$$[\mathbf{G}]_{ij} = \begin{cases} r_i \mathbb{1}_H(i) = \frac{e^{a_i\tau}-1}{a_i\tau} \mathbb{1}_H(i), & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases} \tag{24}$$

Note that the non-zero elements in matrices $\mathbf{D}$ and $\mathbf{E}$ are determined by the RC network *topology*, and $\mathbf{G}$ is a diagonal matrix.

Next, we further construct two coefficient matrices:

$$\mathbf{A} = \begin{bmatrix} -\mathbf{E} & \mathbf{D} & & & & \\ & -\mathbf{E} & \mathbf{D} & & & \\ & & \ddots & \ddots & & \\ & & & -\mathbf{E} & \mathbf{D} & \\ & & & & -\mathbf{E} \end{bmatrix} \in \mathbb{R}^{NK \times NK}, \tag{25}$$

$$\mathbf{B} = \begin{bmatrix} -\mathbf{G} & & & & \\ & -\mathbf{G} & & & \\ & & -\mathbf{G} & & \\ & & & \ddots & \\ & & & & -\mathbf{G} \end{bmatrix} \in \mathbb{R}^{NK \times NK}. \tag{26}$$

Note that A is in a *blockwise Jordan normal form* [27] and $\mathbf{B}$ is diagonal. Lastly, we let $\mathbf{T}_A \triangleq [T_A[1](\mathbf{s} + \text{Diag}\{\mathbf{r}\}\mathbf{o}[1])^\top \ \dots \ T_A[K](\mathbf{s} + \text{Diag}\{\mathbf{r}\}\mathbf{o}[K])^\top]^\top \in \mathbb{R}^{NK}$. With the above definitions of matrices and vectors, we can rewrite the linearized RC thermal dynamics in (19) as $\mathbf{AT}+\mathbf{Bu} = \mathbf{T}_A$. Further, we can compactly rewrite Problem R-HPrO in matrix form as follows:

R-HPrO-M:

$$\begin{aligned} \underset{\mathbf{T},\mathbf{u}}{\text{Minimize}} \quad & \sum_{k=1}^{K}\sum_{i=1}^{N} p[k]u_i[k] + P_{\text{peak}}(z) \\ \text{subject to} \quad & \mathbf{1}^\top \mathbf{u}[k] - z \leq 0, \ \forall k \in \Gamma, && (27) \\ & \mathbf{AT} + \mathbf{Bu} = \mathbf{T}_A && (28) \\ & \mathbf{T}^{lb} \leq \mathbf{T} \leq \mathbf{T}^{ub}, && (29) \\ & \mathbf{0} \leq \mathbf{u} \leq \mathbf{u}^{\max}, && (30) \end{aligned}$$

where $\mathbf{T}^{lb} \triangleq [T_i^{lb}, \forall i]^\top \in \mathbb{R}^{NK}$, $\mathbf{T}^{ub} \triangleq [T_i^{ub}, \forall i]^\top \in \mathbb{R}^{NK}$, and $\mathbf{u}^{\max} \triangleq u^{\max}\mathbf{1} \in \mathbb{R}^{NK}$. In Problem R-HPrO-M, the inequalities in (29) and (30) are entry-wise.

We note that the objective function in Problem R-HPrO-M is separable and the reposed RC thermal dynamics constraint in (29) can also be separated block-wise in terms of variables $\mathbf{T}$ and $\mathbf{u}$. Further since Problem R-HPrO is convex and it is not difficult to check that the Slater's condition holds, we can conclude that the strong duality holds and we can solve Problem R-HPrO in its dual domain. This prompts us to develop an ADMM-type [28] *dual decomposition* scheme. Toward this end, we associate dual variables $\boldsymbol{\mu} \in \mathbb{R}_+^{|\mathbf{\Gamma}|}$ with constraint (27) and $\mathbf{v} \in \mathbb{R}^{NK}$ with (28), respectively. Let $\rho > 0$ be some fixed constant chosen before running the algorithm. Then, we can formulate a ρ-parameterized *augmented Lagrangian* as follows:

$$\begin{aligned} L_\rho(\mathbf{T}, \mathbf{u}) \triangleq & \sum_{k=1}^{K}\sum_{i=1}^{N} p[k]u_i[k] + \hat{P}(z) + \\ & \sum_{k\in\Gamma} \mu_k(z - \mathbf{1}^\top \mathbf{u}[k]) + \mathbf{v}^\top(\mathbf{AT}+\mathbf{Bu}-\mathbf{T}_A) \\ & + \frac{\rho}{2}\Big(\|\mathbf{AT}+\mathbf{Bu}-\mathbf{T}_A\|_2^2 + \sum_{k\in\Gamma}(z-\mathbf{1}^\top\mathbf{u}[k])^2\Big). \end{aligned} \tag{31}$$

Based on the augmented Lagrangian in (31), we can derive the primal and dual updating schemes for Problem R-HPrO as follows:

a) Primal Updates: We first derive the updates for the primal temperature setpoints $\mathbf{T}[k]$, $k = 1, \dots, K$, for which we have the following result:

Proposition 2 (Temperature setpoints update)**.** *The primal temperature setpoints $\mathbf{T}[k]$, $k = 1, \dots, K$, can be computed as follows:*

$$\mathbf{T}[k] = \begin{cases} \frac{1}{\rho}\mathbf{E}^{-1}(-\mathbf{E}^\top \mathbf{v}[1] + \rho\mathbf{DT}[2]), & k = 1 \\ \frac{1}{\rho}\mathbf{E}^{-1}(\mathbf{D}^\top \mathbf{v}[k-1]) - \mathbf{E}^\top \mathbf{v}[k] + \rho\mathbf{DT}[k+1]), \\ \qquad k = 2, \dots, K-1, \\ \frac{1}{\rho}\mathbf{E}^{-1}(\mathbf{D}^\top \mathbf{v}[K-1]) - \mathbf{E}^\top \mathbf{v}[K]), & k = K. \end{cases} \tag{32}$$

Proof. Take partial derivative of $L_\rho(\mathbf{T}, \mathbf{u})$ with respect to $\mathbf{T}[k]$ and set it to 0. Then, by further using the blockwise Jordan normal form structure of $\mathbf{A}$ in (25), we can obtain that:

$$\frac{\partial L_\rho(\mathbf{T}, \mathbf{u})}{\partial \mathbf{T}[k]} = \begin{cases} -\mathbf{E}^\top \mathbf{v}[1] + \rho(-\mathbf{ET}[1] + \\ \quad \mathbf{DT}[2]) = \mathbf{0}, \ k = 1, \\ \mathbf{D}\top \mathbf{v}[k-1] - \mathbf{E}^\top \mathbf{v}[k] + \rho(-\mathbf{ET}[k] + \\ \quad \mathbf{DT}[k+1]) = \mathbf{0}, \ k = 2, \dots, K-1, \\ \mathbf{D}^\top \mathbf{v}[K-1] - \mathbf{E}^\top \mathbf{v}[K] + \\ \quad \rho(-E\mathbf{T}[K]) = \mathbf{0}, k = K. \end{cases} \tag{33}$$

Then, the results stated in Proposition 2 follow from solving for $\mathbf{T}[K]$ from the three cases in (33) correspondingly. This completes the proof. □

Remark 1. It is important to note that the structural property of (32) implies that $\mathbf{T}[k]$, $k = 1, \dots, K$, can be efficiently computed in a *backward induction* fashion: Starting from $k = K$ and using the third equation in (32), each $\mathbf{T}[k-1]$ can be computed by $\mathbf{T}[k]$ using the second equation in (32),

and this process will continue until $k = 1$, for which the result can be computed using the first equation in (32).

Next, we derive the updates for the primal HVAC control decisions $\mathbf{u}[k]$, $k = 1, \ldots, K$, for which we have the following result:

Proposition 3 (HVAC Control Decision)**.** *The primal HVAC decisions $\mathbf{u}[k]$, $k = 1, \ldots, K$, can be computed as follows:*

$$\mathbf{u}[k] = \frac{1}{\rho}\big[\mathbf{G} + \mathbb{1}_\Gamma(k)\mathbf{1}\mathbf{1}^\top\big]^{-1}\big[[p[k] - (\mu_k + \rho z) \mathbb{1}_\Gamma(k)]\mathbf{1} - \mathbf{G}^\top \mathbf{v}[k]\big]. \quad (34)$$

Further, each entry in the matrix inversion $\big[\mathbf{G} + \mathbb{1}_\Gamma(k)\mathbf{1}\mathbf{1}^\top\big]^{-1}$ in (34) can be computed efficiently in closed-form as follows:

$$\big[\mathbf{G} + \mathbb{1}_\Gamma(k)\mathbf{1}\mathbf{1}^\top\big]^{-1}_{ij} = \begin{cases} \frac{1 + \mathbb{1}_\Gamma(k)\sum_{j'=1,\neq i}^{N}(1/r_j)}{r_i(1 + \mathbb{1}_\Gamma(k)\sum_{j=1}^{N}(1/r_{j'}))}, & i = j, \\ \frac{(1/r_j)\mathbb{1}_\Gamma(k)}{r_i(1 + \mathbb{1}_\Gamma(k)\sum_{j=1}^{N}(1/r_{j'}))}, & i \neq j. \end{cases} \quad (35)$$

Proof. By taking the partial derivative of $L_\rho(\mathbf{T}, \mathbf{u})$ with respect to $\mathbf{u}[k]$ and setting it to 0, and further using the special structure of $\mathbf{B}$, one can obtain that:

$$\frac{\partial L_\rho(\mathbf{T}, \mathbf{u})}{\partial \mathbf{u}[k]} = p[k]\mathbf{1} - \mu_k \mathbb{1}_\Gamma(k)\mathbf{1} - \mathbf{G}^\top \mathbf{v}[k] - \rho \mathbf{G}\mathbf{u}[k] - \mathbb{1}_\Gamma(k)[(z - \mathbf{1}^\top \mathbf{u}[k])\mathbf{1}] = \mathbf{0}, \ \forall k, \quad (36)$$

where $\mathbb{1}_\Gamma(k)$ is an indicator function that takes value 1 if $k \in \Gamma$ and 0 otherwise. It then follows from (36) that:

$$\rho\big[\mathbf{G} + \mathbb{1}_\Gamma(k)\mathbf{1}\mathbf{1}^\top\big]\mathbf{u}[k] = [p[k] - (\mu_k + \rho z)\mathbb{1}_\Gamma(k)]\mathbf{1} - \mathbf{G}^\top \mathbf{v}[k]. \quad (37)$$

Solving for $\mathbf{u}[k]$ from (37) yields the result stated in (34).

Next, noting from the definition in (23) that $\mathbf{G}$ is diagonal, we have that $\mathbf{G}^{-1}$ is also diagonal and $[\mathbf{G}^{-1}]_{ii} = 1/r_i$. Now, it is important to recognize that $\mathbf{G} + \mathbb{1}_\Gamma(k)\mathbf{1}\mathbf{1}^\top$ is a *rank-1* update to $\mathbf{G}$. Therefore, $\big[\mathbf{G} + \mathbb{1}_\Gamma(k)\mathbf{1}\mathbf{1}^\top\big]^{-1}$ can be computed by using the *Sherman-Morrison-Woodbury* (SMW) matrix inversion lemma [27] as follows:

$$\big[\mathbf{G} + \mathbb{1}_\Gamma(k)\mathbf{1}\mathbf{1}^\top\big]^{-1} = \mathbf{G}^{-1} - \frac{\mathbf{G}^{-1}\mathbf{1}\mathbf{1}^\top\mathbf{G}^{-1}}{1 + \mathbb{1}_\Gamma(k)\mathbf{1}^\top\mathbf{G}^{-1}\mathbf{1}}\mathbb{1}_\Gamma(k). \quad (38)$$

Plugging in the definition of $\mathbf{G}$ (cf. Eq. (24)) into (38) yields the result in (35). This completes the proof. □

Remark 2. From Proposition 3, we can see that if time-slot $k \notin \Gamma$, i.e., in the off-peak time period, then $\mathbf{u}[k]$ can be computed distributively using local information since $\mathbf{G}$ is block diagonal. On the other hand, if time-slot $k \in \Gamma$, i.e., in the peak-time period, then the $u[k]$ solution can still be computed in a distributed fashion by exchanging the r_i-information between each zone.

Lastly, taking the derivative of $L_\rho(\mathbf{T}, \mathbf{u})$ with respect to z and setting it to 0, we obtain $\frac{\partial L_\rho(\mathbf{T}, \mathbf{u})}{\partial z} = \sum_{k\in\Gamma}\mu_k + \frac{\hat{P}(z)}{dz} + \sum_{k\in\Gamma}\rho(z - \mathbf{1}^\top\mathbf{u}[k]) = 0$, which further leads to the following primal z-solution:

$$z = \frac{1}{\rho|\Gamma|}\bigg[\sum_{k\in\Gamma}\rho\mathbf{1}^\top\mathbf{u}[k] - \sum_{k\in\Gamma}\mu_k - \frac{d\hat{P}(z)}{dz}\bigg]. \quad (39)$$

2) Dual update: For notational simplicity, we let $\mathbf{v}^+$ and $\boldsymbol{\mu}^+$ represent the values of dual variables $\mathbf{v}$ and $\boldsymbol{\mu}$ in the next iteration, respectively. Then, the dual variable updates can be written as follows:

$$\mathbf{v}^+ = \mathbf{v} + \rho\big[(\mathbf{A}\mathbf{T} + \mathbf{B}\mathbf{u} - \mathbf{T}_A)\big] + \begin{bmatrix} (z - \mathbf{1}^\top\mathbf{u}[1])\mathbb{1}_\Gamma(1) \\ \vdots \\ (z - \mathbf{1}^\top\mathbf{u}[K])\mathbb{1}_\Gamma(K) \end{bmatrix}, \quad (40)$$

$$\mu_k^+ = \mu_k + \rho(z - \mathbf{1}^\top\mathbf{u}[k]). \quad (41)$$

Finally, by combining the primal and dual updates, we have the following algorithm:

Algorithm 1: An Efficient Dual-Based Distributed Approach for Solving Problem R-HPrO-M.

Initialization:

1. For each thermal capacitor and resistor, choose some appropriate initial values for temperature setpoints $\mathbf{T}[k]$ and HVAC control decisions $\mathbf{u}[k]$, $\forall k = 1, \ldots, K$.
2. For each thermal capacitor i, choose appropriate initial values for dual variables $\mathbf{v}[k]$, $k = 1, \ldots, K$. For the central HVAC controller, choose appropriate initial values for z.

Main Loop:

3. *Primal Temperature Setpoints Update:* Based on (32), compute and update the temperature setpoints $\mathbf{T}[k]$, $k = 1, \ldots, K$, in a backward induction fashion.
4. *Primal HVAC Control Decisions Update:* Based on (34) and (35), compute and update the HVAC control decisions $\mathbf{u}[k]$, $k = 1, \ldots, K$ either distributively or in an SMW fashion (depending on whether or not time-slot k is in the peak demand period Γ). Use the computed $\mathbf{u}[k]$-information to update z following (39).
5. *Dual Variable Updates:* Update the dual variables $\mathbf{v}[k]$ and μ_k, $k = 1, \ldots, K$, according to (40) and (41). Let $t = t+1$.
6. Terminate the algorithm if the algorithm converges or if a predefined run-time limit is reached. Otherwise, go back to Step 3 and repeat the whole primal and dual update processes.

In Algorithm 1, after initializing in Steps 1–2, Steps 3–4 are for primal updates, while Step 5 is for dual updates, respectively. The main iteration stops if the criterion in Step 6 is met. The convergence of the proposed algorithm follows similarly from that of the ADDM approach and thus is omitted for brevity in this paper.

Time-Complexity Analysis: The time-complexity of the proposed distributed algorithm is stated as follows.

Proposition 4. *The time-complexity of the dual-based distributed algorithm in Algorithm 1 is $O(KN^3\log(1/\epsilon))$, where ϵ represents some desired accuracy for stopping criterion.*

Proof. In the main loop in Algorithm 1, Step 3 is based on Eq. (32), where the complexity is dominated by computing $\mathbf{E}^{-1}$, which is of $O(KN^3)$. In Step 4, Eqs. (34) and (35) use the SMW matrix inversion lemma, which is of $O(KN^2)$. In Step 5, Eqs. (40) and (41) are simple linear transformation, which are of $O(KN)$. Combining Steps 3–5, it can be seen that the per-iteration time-complexity in the main loop is

$O(KN)$ Lastly, due to the strong convexity assumption of $\hat{P}(\cdot)$ and the ADMM-type structure, it follows from [29] that the algorithm has a linear convergence rate in terms of the main loop iterations, which further implies the stated time-complexity result. This completes the proof. □

Proposition 4 shows that, by exploiting special problem structure, our custom-designed distributed algorithm has a lower time-complexity than the $O(K^3N^3\log(1/\epsilon))$ time-complexity of the interior-point method, which is a general method for solving convex optimization problems.

Memory-Complexity Analysis: Compared to the centralized method, the proposed distributed method is also advantageous in terms of memory complexity. Specifically, thanks to the ADMM-type decomposition structure, each zone i in the RC-network only needs to store $|\mathcal{N}_{i0}|$ p-values and $|\mathcal{N}_{i0}|$ q-values in the primal temperature setpoints update (see Step 3 in Algorithm 1); and single N values of $\{[\mathbf{G}+\mathbb{1}_\Gamma(k)\mathbf{1}\mathbf{1}^\top]^{-1}_{ij}\}_{j=1}^N$ in the primal HVAC control decision update (see Step 4 in Algorithm 1). Similarly, from Step 5 in Algorithm 1, we have that each node i needs to store $|\mathcal{N}_{i0}|$ p-values and $|\mathcal{N}_{i0}|$ q-values for dual variable updates. By contrast, in a centralized implementation, we need to store the $\mathbf{A}$ and $\mathbf{B}$ matrices, all of which are of size N^2K^2. In other words, in centralized implementation, one needs a node that has large memory to hold all problem-related parameters, which is problematic when N and K are large. On the other hand, in our distributed implementation, the problem-related parameters are distributed evenly at each node, and there is no need to have a "super node" that has a large memory requirement to store all problem-related information.

6 Numerical Results

In this section, we perform numerical experiments to validate our theoretical results. We use the RC thermal transfer model to simulate multi-zone buildings and investigate the effects of different building parameters, internal loads, and external environments on the cooling energy cost reduction by comparing with some existing strategies. The weather data used in simulation come from the third typical meteorological year collection (TMY3) in Columbus, Ohio; Minneapolis, Minnesota; Baltimore, Maryland; Houston, Texas; and Los Angeles (LA), California. Base on the [30], the total thermal resistance and heat capacitance of the external walls for the prototype multi-zone building model in Columbus and Minneapolis are set to 2.97 $m^2\cdot$ K/W and 134.80 kJ/m^2·K, respectively. And in the other three cities mentioned above, the thermal resistance and heat capacitance of the external walls are 2.16 $m^2\cdot$ K/W and 49.39 kJ/m^2·K, respectively. The other parameters for the prototype multi-zone building model are kept the same for these five cities during simulation. For example, the windows in external walls have a u factor of 5.78 W/$m^2\cdot$ K; the window-to-wall ratio is 50%; and the interior partition walls consist of one layer of gypsum board on each side, the air gap, and batt insulation, of which the total thermal resistance and heat capacitance are 2.57 m^2·K/W and 13.86 kJ/$m^2\cdot$ K, respectively. Based on these assembly types, parameter values of the RC network model are determined. The occupancy schedule and plug loads (0.04 kW/m^2 during the occupied time and 8.00×10^{-4} kW/m^2 during unoccupied time) are primarily based on the schedules from an academic building on the campus. The occupied time is from 7 AM to 11 PM, and the peak time is from 12 PM to 6 PM.

TABLE 1
Setpoint temperatures of baseline strategies.

Strategies	11 PM to 7 AM	7 AM to 12 PM	12 PM to 6 PM	6 PM to 11 PM
ON/OFF	-	22.5 °C	22.5 °C	24.5 °C
PC+LU	24.5 °C	21.8 °C	Linear-up from 21.8 °C to 24.5 °C	24.5 °C
PC+CI	24.5 °C	21.8 °C	Raising-up based on (43) from 21.8 °C to 24.5 °C	24.5 °C
EPC+LU	24.5 °C before 4 AM and 21.8 °C after 4 AM	21.8 °C	Linear-up from 21.8 °C to 24.5 °C	24.5 °C
EPC+CI	24.5 °C before 4 AM and 21.8 °C after 4 AM	21.8 °C	Raising-up based on (43) from 21.8 °C to 24.5 °C	24.5 °C

In order to maintain indoor thermal comfort, the index Predicted Mean Vote (PMV) is always kept between -0.5 and $+0.5$ (equivalent to PPD = 10%) during the occupied time. Based on the CBE Thermal Comfort Tool [31], the lower and upper bounds of setpoint temperature during the occupied time are 21.8 °C and 24.5 °C, corresponding to PMV= -0.5 and PMV= $+0.5$, respectively. The lower and upper bounds of setpoint temperature during unoccupied time are 19.0 °C and 25 °C, respectively. The electricity rates are based on the rate schedules in the power rate zone served by the American Electric Power (AEP). Based on the time-of-day (TOD) schedule, the energy charge is 2.27 cents per kWH from 7 AM to 9 PM local time for all weekdays, and 0.04 cents per kWH from 9 PM to 7 AM for all weekdays, and all hours of the day on Saturdays and Sundays. The demand charge is $4.16 per kW in each month.

The performance of our proposed optimal strategy based on the RC-network is compared with five baseline cases, i.e., occupancy-driven ON/OFF (ON/OFF), precooling combined with linear-up (PC+LU), precooling combined with concave-increasing (PC+CI), extended precooling combined with linear-up (EPC+LU), and extended precooling combined with concave-increasing (EPC+CI), and measured as the cost reduction ratio, which is shown in (42). The detailed setpoint temperatures of the five baseline strategies are shown in Table 1.

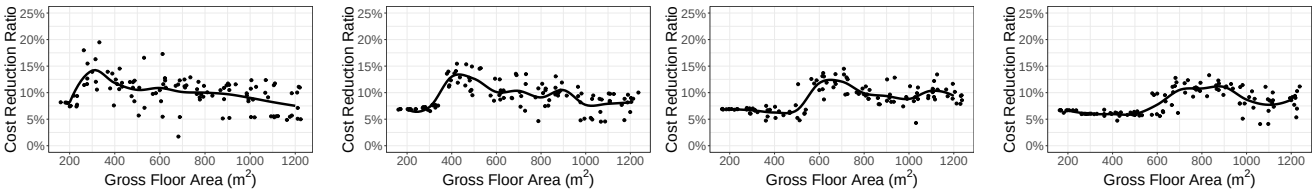
$$\text{Cost Reduction Ratio}=\frac{\text{Cost}_{\text{Base}}-\text{Cost}_{\text{RC-Network}}}{\text{Cost}_{\text{Base}}}\times100\%. \quad (42)$$

$$T[k]=2.71\times\left[1-e^{-(k-12)}\right]+21.8, k\in\Gamma, \quad (43)$$

where Γ denotes the set of time-slots in the peak period, k represents any given time during the peak time, and $T[k]$ is the room setpoint temperature at the time of k.

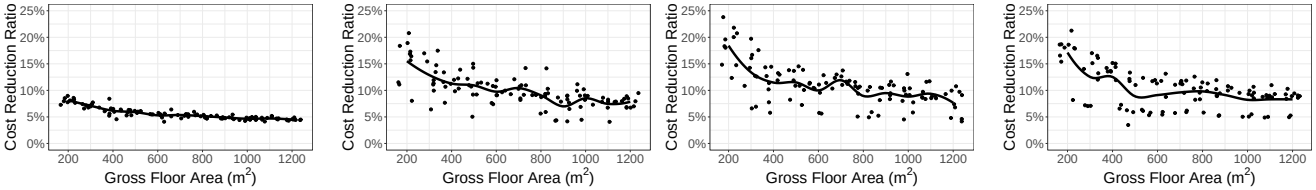
6.1 Building Parameters

In this section, we examine the effects of five building parameters on the performance of the optimal strategy based



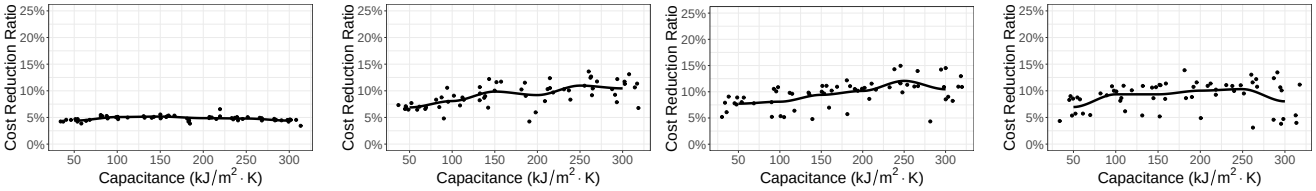
(a) Average room size 40 m^2 (b) Average room size 60 m^2 (c) Average room size 80 m^2 (d) Average room size 100 m^2

Fig. 5. Cost reduction ratio with respect to average room size and gross floor area.



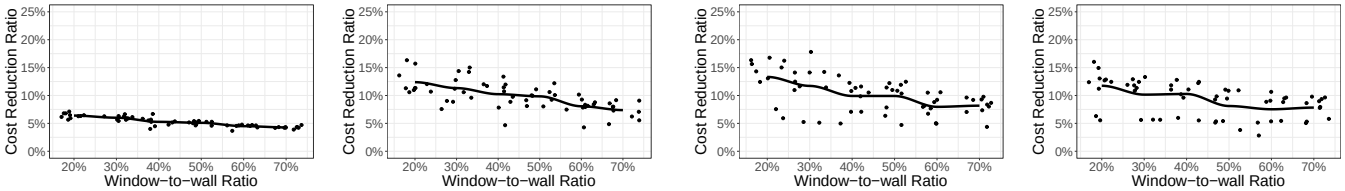
(a) 5 Rooms (b) 10 Rooms (c) 15 Rooms (d) 20 Rooms

Fig. 6. Cost reduction ratio with respect to total number of rooms and gross floor area.



(a) 5 Rooms (b) 10 Rooms (c) 15 Rooms (d) 20 Rooms

Fig. 7. Cost reduction ratio with respect to total number of rooms and wall capacitance.



(a) 5 Rooms (b) 10 Rooms (c) 15 Rooms (d) 20 Rooms

Fig. 8. Cost reduction ratio with total number of rooms and window-to-wall ratio.

on the cooling energy cost reduction ratio. These parameters include i) the gross floor area, ii) average room size, iii) total number of rooms, iv) wall capacitance, and v) the window-to-wall ratio. Table 2 shows the values assigned for each of these building parameters. For each combination of parameter values, we randomly generate 10 building samples for daily simulation. The total number of samples is 1,640. The weather data (one of the hottest summer days) and the external walls for the prototype multi-zone building model in Columbus are used in this section. Our simulation results show that our optimization algorithm achieves approximately 5-20% energy cost reduction compared with the baseline case of ON/OFF strategy with full internal loads, depending on building parameter values.

The effects of the gross floor area in conjunction with average room size on the cost reduction ratio are illustrated in Figure 5. With a fixed average room size (e.g., 40 m^2), buildings with larger gross floor areas will have larger total thermal capacitance (i.e., thermal mass). But the capacitance per unit volume of these buildings would be quite similar. At any given gross floor area (e.g., 800 m^2), buildings with a larger average room size have less internal walls, which leads to smaller capacitance per unit volume for these buildings. Generally, buildings with larger thermal capacitance per unit volume will be able to shift more cooling loads

TABLE 2
Building parameters tested in simulation models.

Parameters	Values
Gross Floor Area	Every 100 m^2 from 200 m^2 to 1,200 m^2
Average Room Size	{40, 60, 80, 100} m^2
Total Number of Rooms	5, 10, 15, 20
Capacitance of External Walls	Every 50 kJ/m^2·K from 50 to 300 kJ/m^2·K
Window-to-wall Ratio	{20%, 30%, 40%, 50%, 60%, 70%}

from the peak time to the off-peak time. As shown in Figure 5, when the average room size is fixed, the trend line of the cost reduction ratio in each subfigure stays flat, which means that the gross floor area in this case has very little effect on the cost reduction ratio. When we compare the trend lines across these four plots, we found that the cost reduction ratios decrease gradually with increasing average room sizes where capacitance per unit volume is decreasing. These findings are consistent with the assumptions.

Figure 6 shows the effects of the gross floor area and the total number of rooms on the cost reduction ratio. When the total number of rooms is fixed (e.g., 5 rooms), buildings with larger gross floor area will have smaller thermal capacitance per unit volume. At any given gross floor area (e.g., 600 m^2), buildings having a larger total number of rooms have more internal walls, which result in larger thermal capacitance per unit volume. As seen in Figure 6, in general, the cost reduction ratio decreases with the increasing gross floor area when the total number of rooms is fixed. When the total number of rooms varies, buildings with smaller total number of rooms incur lower cost reduction ratios. The ratios in Figure 6(a) are apparently lower than that in the other three plots due to the much lower thermal capacitance per unit volume.

The effects of external wall capacitance and the number of rooms on the cost reduction ratio are shown in Figure 7 with a given gross floor area (using 800 m^2 in this study). In general, with a fixed total number of rooms, the cost reduction ratio increases gradually when the capacitance of external walls increases from 50 $kJ/m^2 \cdot K$ (representing lightweight building construction without much thermal mass in the walls) to 300 $kJ/m^2 \cdot K$ (representing heavy-weight construction with large amounts of thermal mass materials incorporated). With the increasing total number of rooms, the trend line of cost reduction ratio is shifted upward along the vertical axis. Similarly, the trend line for 5 rooms is apparently lower than the other three trend lines.

Figure 8 illustrates the effects of the window-to-wall ratio and the total number of rooms on the cost reduction ratio. In general, the cost reduction ratio decreases with the increasing window-to-wall ratio. Because the capacitance of the window is much lower than the capacitance of the external wall, a larger window-to-wall ratio would lead to a smaller capacitance per unit volume for the building, which reduces the building's ability to store energy and shift HVAC loads. The observation across four plots is consistent with our early observations.

Lastly, two typical examples of a room setpoint temperature trajectory in a 24-hour period from the optimization process are shown in Figure 9. The majority of building samples with 5 rooms have the setpoint temperature schedules similar to Trajectory 1, while other samples are similar to Trajectory 2.

6.2 Internal Loads

In this section, we examine the effects of the internal loads on the performance of the optimal strategy compared with the baseline case of ON/OFF strategy. The building samples with 20 rooms are simulated with different internal load ratios (i.e., 0.125, 0.5, 0.75, and 1). The internal load ratios

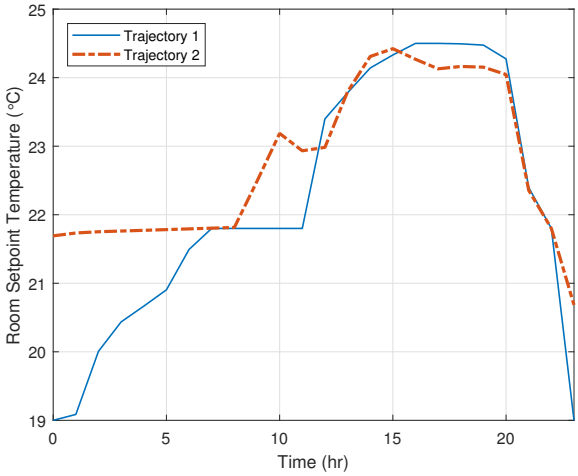
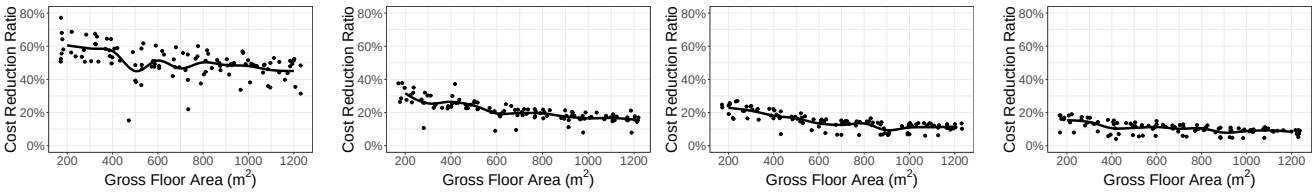


Fig. 9. The room setpoint temperature trajectories.

less than 1 represent buildings with less plug loads or at various occupancy levels. As shown in Figure 10, generally the cost reduction ratio decreases with the increasing internal loads, showing that the effect of precooling on peak load and energy cost reduction would be lower when facing higher internal loads. In [32], all the simulations were based on building samples with much lower internal loads, i.e., approximately 12.5% of the full loads used in this paper. Therefore, the cost reduction ratios in [32] were much higher.

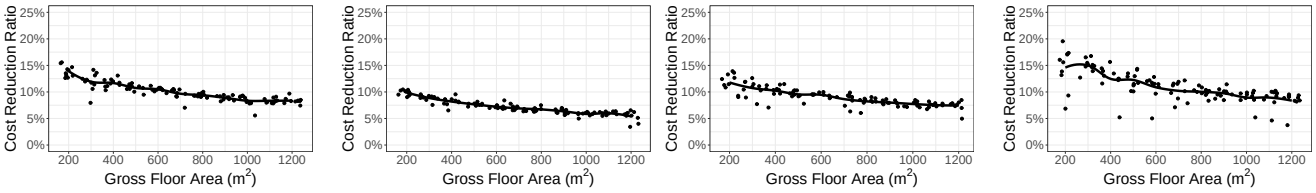
6.3 External Environments

In this section, we examine the effects of the external environments on the performance of the optimal strategy against the ON/OFF baseline case. Besides Columbus, which has been simulated earlier in this paper, the building samples with 20 rooms are simulated in the other four cities, and the results are shown in Figure 11. Based on the ASHRAE Climate Zone definition, these five cities (Houston, LA, Baltimore, Columbus, and Minneapolis) are in climate zones 2A, 3B, 4A, 5A, and 6A, corresponding to the Hot-Humid, Warm-Dry, Mixed-Humid, Cool-Humid, and Cold-Humid zones, respectively. Among these five cities, LA has the lowest temperature in summer. The selected simulation day has the highest temperature of 26.11 °C during the day and the lowest temperature of 16.67 °C during the night. In this weather condition, the advantage of precooling is not obvious, because the building does not require too much cooling during the day. Thus, LA has the lowest cost reduction ratio of 5-10%. In contrast, Houston and Baltimore have the highest summer temperature of these five cities, with the highest temperatures of 33.89 °C and 31.67 °C during the simulation day and the lowest temperatures of 23.33 °C and 22.78 °C during the night, respectively. Compared to the LA building, the buildings in these two cities would consume much more cooling energy during the day. However, the relatively high nighttime temperatures in these locations are not conductive to utilizing precooling. Therefore, Houston and Baltimore only ranked the second highest in the cost reduction ratio with a range of 7.5%-15%. Columbus and Minneapolis have the lower temperatures



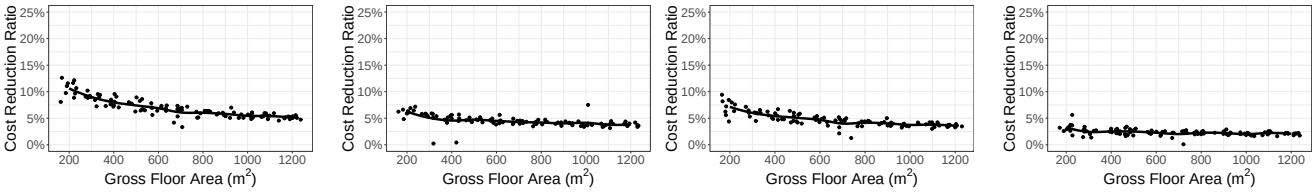
(a) 0.125 internal loads (b) 0.5 internal loads (c) 0.75 internal loads (d) Full internal loads

Fig. 10. Cost reduction ratio with respect to internal loads and gross floor area.



(a) Houston (b) Los Angeles (c) Baltimore (d) Minneapolis

Fig. 11. Cost reduction ratio with respect to external environment and gross floor area.



(a) PC+LU (b) PC+CI (c) EPC+LU (d) EPC+CI

Fig. 12. Optimal strategy compared with baseline cases.

(18.33 °C and 17.78 °C, respectively) than Houston and Baltimore at night and higher temperatures (29.44 °C and 28.33 °C, respectively) than LA during the day, which helps shift more cooling loads from the peak time to the off-peak time, resulting in the highest cost reduction ratio up to 20%.

6.4 Comparison with Other Strategies

In addition to using the ON/OFF strategy as the baseline, this section provides additional comparisons between the proposed optimal strategy and the other four existing strategies from the literature, including PC+LU, PC+CI, EPC+LU, and EPC+CI strategies. The simulation results based on building samples with 20 rooms (located in Columbus) are shown in Figure 12. It can be seen from Figures 6(d) and 12, the optimal strategy outperforms all the selected baselines by achieving positive cost reduction ratios ranging from 2.5% to 20%. Among the five baseline cases, the EPC+CI strategy performs the best with the least cooling energy cost based on the electricity price scheme used in this study. In contrast, the ON/OFF strategy (not utilizing precooling) leads to the highest cooling energy cost.

7 Conclusion

In this paper, we developed an accurate and tractable mathematical framework for multi-zone HVAC precooling optimization, with the goal to minimize total energy costs and peak load demand. The main results of this paper are three-fold: i) We developed an RC-network-based analytical model for multi-zone HVAC precooling to minimize both total energy costs and peak load demand; ii) By exploiting the physical characteristics of HVAC controls, we showed that the HVAC procooling optimization problem based on the proposed RC-network model can be converted into an approximated convex optimization problem, which further leads to efficient optimization algorithm design. Moreover, the convex approximation offers a graceful trade-off between energy cost optimality and problem dimension complexity; and iii) By leveraging special structures in the approximated convex optimization, we designed an efficient ADMM-type distributed algorithm to solve the HVAC precooling optimization problem. Moreover, we showed that low-complexity computation schemes can be developed for temperature setpoints and HVAC control decisions under the ADMM-type algorithmic framework. To verify the efficacy of the proposed analytical models and optimization algorithms, we have conducted extensive simulation studies and investigated the effects of five building parameters, including the gross floor area, average room size, total number of rooms, wall capacitance, and the window-to-wall ratio, on the cooling energy cost reduction ratio based on the weather in Columbus, OH. Overall, when compared with the baseline case of the occupancy-driven ON/OFF strategy, our algorithm was able to achieve approximately 5-20% cooling energy cost reduction, depending on the selected values for various building parameters. We found that with a fixed total number of rooms, the gross floor area and

the window-to-wall ratio had negative effects on the cost reduction ratio; i.e., buildings with larger values of these parameters would generally incur smaller cost reduction ratios. In contrast, buildings with larger values of wall capacitance would result in larger cost reduction ratios. We also examined the effects of internal loads and different weather conditions on the performance of the optimal strategy against the ON/OFF baseline. In general, the cost reduction ratio (ranging from 2.5% to 75%) decreases with the increasing internal loads. Of the five cities (i.e., Houston, LA, Baltimore, Columbus, and Minneapolis) representing different climate zones in the U.S., our algorithm was able to achieve the highest cost reduction ratios (ranging from 7.5% to 20%) in Columbus and Minneapolis and the least cooling energy cost reduction (5-10%) in LA. When compared with five existing strategies (using the ON/OFF, PC+LU, PC+CI, EPC+LU, and EPC+CI strategies as the baselines in simulation), our algorithm outperformed all of them based on the employed electricity price scheme. Collectively, the results and findings in this paper contribute to a new and exciting research paradigm that leverages HVAC precooling optimization to significantly improve environmental sustainability of buildings. Future research topics may include investigating the effects of other parameters (e.g., the electricity rate and occupancy schedule) on the cost reduction ratio, as well as the impacts of their prediction errors.

REFERENCES

[1] W. Wang, R. Zmeureanu, and H. Rivard, "Applying multi-objective genetic algorithms in green building design optimization," *Building and environment*, vol. 40, no. 11, pp. 1512–1525, 2005.

[2] U. Congress, "Office of technology assessment.(1992)," *Building Energy Efficiency*, 2006.

[3] B. E. D. Book, "Us department of energy, 2011," 2010.

[4] Y. Sun, S. Wang, and G. Huang, "A demand limiting strategy for maximizing monthly cost savings of commercial buildings," *Energy and buildings*, vol. 42, no. 11, pp. 2219–2230, 2010.

[5] J. E. Seem, "Adaptive demand limiting control using load shedding," *HVAC&R Research*, vol. 1, no. 1, pp. 21–34, 1995.

[6] J. J. H. . Associates. (2018) The Home of DOE-2 Based Building Energy Use and Cost Analysis Software. [Online]. Available: http://doe2.com/

[7] T. Salsbury, P. Mhaskar, and S. J. Qin, "Predictive control methods to improve energy efficiency and reduce demand in buildings," *Computers & Chemical Engineering*, vol. 51, pp. 77–85, 2013.

[8] M. F. Haniff, H. Selamat, R. Yusof, S. Buyamin, and F. S. Ismail, "Review of hvac scheduling techniques for buildings towards energy-efficient and cost-effective operations," *Renewable and Sustainable Energy Reviews*, vol. 27, pp. 94–103, 2013.

[9] R. Yin, P. Xu, M. A. Piette, and S. Kiliccote, "Study on auto-dr and pre-cooling of commercial buildings with thermal mass in california," *Energy and Buildings*, vol. 42, no. 7, pp. 967–975, 2010.

[10] M. Arnold and G. Andersson, "Model predictive control of energy storage including uncertain forecasts," in *Proc. Power Systems Computation Conference (PSCC), Stockholm, Sweden*, vol. 23, 2011, pp. 24–29.

[11] S. Privara, J. Siroky, L. Ferkl, and J. Cigler, "Model predictive control of a building heating system: The first experience," *Energy and Buildings*, vol. 43, no. 2, pp. 564–572, 2011.

[12] A. Rabl and L. Norford, "Peak load reduction by preconditioning buildings at night," *International Journal of Energy Research*, vol. 15, no. 9, pp. 781–798, 1991.

[13] F. Morris, J. E. Braun, and S. Treado, "Experimental and simulated performance of optimal control of building thermal storage," *Ashrae transactions*, vol. 100, no. 1, pp. 402–414, 1994.

[14] K. R. Keeney and J. E. Braun, "Application of building precooling to reduce peak cooling requirements," *ASHRAE transactions*, vol. 103, no. 1, pp. 463–469, 1997.

[15] J. Ma, J. Qin, T. Salsbury, and P. Xu, "Demand reduction in building energy systems based on economic model predictive control," *Chemical Engineering Science*, vol. 67, no. 1, pp. 92–100, 2012.

[16] J. Siroky, F. Oldewurtel, J. Cigler, and S. Privara, "Experimental analysis of model predictive control for an energy efficient building heating system," *Applied energy*, vol. 88, no. 9, pp. 3079–3087, 2011.

[17] F. Oldewurtel, A. Parisio, C. N. Jones, M. Morari, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and K. Wirth, "Energy efficient building climate control using stochastic model predictive control and weather predictions," in *Proc. American control conference (ACC)*. IEEE, 2010, pp. 5100–5105.

[18] K.-h. Lee and J. E. Braun, "Development of methods for determining demand-limiting setpoint trajectories in buildings using short-term measurements," *Building and Environment*, vol. 43, no. 10, pp. 1755–1768, 2008.

[19] ——, "Model-based demand-limiting control of building thermal mass," *Building and Environment*, vol. 43, no. 10, pp. 1633–1646, 2008.

[20] S. Mukherjee, S. Mishra, and J. T. Wen, "Building temperature control: A passivity-based approach," in *Decision and Control (CDC), 2012 Proc. IEEE 51st Annual Conference on*. IEEE, 2012, pp. 6902–6907.

[21] S. K. Gupta, K. Kar, S. Mishra, and J. T. Wen, "Collaborative energy and thermal comfort management through distributed consensus algorithms," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 4, pp. 1285 – 1296, October 2015.

[22] S. Bhattacharya, K. Kar, and J. H. Chow, "Optimal precooling of thermostatic loads under time-varying electricity prices," in *Proc. IEEE American Control Conference (ACC)*. IEEE, 2017, pp. 1407–1412.

[23] S. Wang and X. Xu, "Parameter estimation of internal thermal mass of building dynamic models using genetic algorithm," *Energy conversion and management*, vol. 47, no. 13, pp. 1927–1941, 2006.

[24] S. F. Fux, A. Ashouri, M. J. Benz, and L. Guzzella, "Ekf based self-adaptive thermal model for a passive house," *Energy and Buildings*, vol. 68, pp. 811–817, 2014.

[25] T. Weber, G. Johannesson, M. Koschenz, B. Lehmann, and T. Baumgartner, "Validation of a fem-program (frequency-domain) and a simplified rc-model (time-domain) for thermally activated building component systems (tabs) using measurement data," *Energy and Buildings*, vol. 37, no. 7, pp. 707–724, 2005.

[26] X. Xu and S. Wang, "A simplified dynamic model for existing buildings using ctf and thermal network models," *International Journal of Thermal Sciences*, vol. 47, no. 9, pp. 1249–1262, 2008.

[27] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York, NY: Cambridge University Press, 1990.

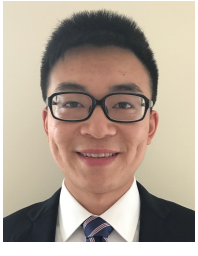
[28] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1 – 122, 2011.

[29] R. Nishihara, L. Lessard, B. Recht, A. Packard, and M. I. Jordan, "A general analysis of the convergence of ADMM," in *Proc. of ICML*, Lille, France, 2015.

[30] M. Deru, K. Field, D. Studer, K. Benne, B. Griffith, P. Torcellini, B. Liu, M. Halverson, D. Winiarski, M. Rosenberg, M. Yazdanian, J. Huang, and D. Crawley, "Us department of energy commercial reference building models of the national building stock," 2011.

[31] T. Hoyt, S. Schiavon, A. Piccioli, D. Moon, and K. Steinfeld, "Cbe thermal comfort tool," *Center for the Built Environment, University of California Berkeley*, 2013.

[32] H. Shi, J. Liu, and Q. Chen, "Hvac precooling optimization for green buildings: An rc-network approach," in *Proceedings of the Ninth International Conference on Future Energy Systems*. ACM, 2018, pp. 249–260.



Hongsen Shi received a Ph.D. degree in Agriculture Engineering with specialization in construction systems managment, an M.S. degree in Civil Engineering, and an M.S. degree in Statistics at the Ohio State University in 2015 and 2018, respectively. His research focuses on building energy efficiency improvement, building thermal comfort diagnosis, and HVAC precooling optimization.

Jia Liu (S'03–M'10–SM'16) received his Ph.D. degree in the Bradley Department of Electrical and Computer Engineering at Virginia Tech, Blacksburg, VA in 2010. Since April 2010, he has been working as a Postdoctoral Researcher in the Department of Electrical and Computer Engineering at The Ohio State University. He is currently an Assistant Professor in the Dept. of Computer Science at Iowa State University, where he joined in Aug. 2017. His research areas include theoretical foundations of control and optimization for stochastic networked systems, distributed algorithms design, optimization of cyber-physical systems, data analytics infrastructure, and machine learning. Dr. Liu is a senior member of IEEE, a member of ACM, and a member of SIAM. His work has received numerous awards at top venues, including IEEE INFOCOM'19 Best Paper Award, IEEE INFOCOM'16 Best Paper Award, IEEE INFOCOM'13 Best Paper Runner-up Award, IEEE INFOCOM'11 Best Paper Runner-up Award, and IEEE ICC'08 Best Paper Award. He is a recipient of Bell Labs President Gold Award in 2001. He has served as a TPC member for IEEE INFOCOM since 2010 and a TPC member of MobiHoc since 2017. His research has been supported by NSF, AFOSR, AFRL, and ONR.

Qian Chen received her B.E. degree in Architecture from Southwest Jiaotong University in China in 1996 and an M.E. degree in Civil Engineering from National University of Singapore in 2003. She received a second M.S. degree in Architecture in 2004 and her Ph.D. degree in Environmental Design and Planning/Building Construction in 2007, both from Virginia Tech, Blacksburg, VA. She is currently an Associate Professor of Construction Systems Management at The Ohio State University. Before coming to the U.S., she had six years of construction industry work experience. Dr. Chen's research focuses on Interface Management in Integrated Project Delivery, Building Information Modeling, IT Applications in Construction Management, Lean Construction, Occupational Safety and Health, and High Performance Green Buildings. She is a LEED Accredited Professional and serves as the Editor for Journal of Green Building.

APPENDIX

Since the the time-slot duration τ is sufficiently short, we have $u_i^{[k]}(t) = u_i[k]/\tau$, $o_i^{[k]}(t) = o_i[k]/\tau\ \forall i,k,t$, and the temperature evolution curve of each zone i is a line segment connecting $T_i[k]$ and $T_i[k+1]$, we can rewrite the RC thermal dynamics ODE as follows:

$$C_i\frac{dS_i^{[k]}(t)}{dt}=\sum_{j\in\mathcal{N}_i}\frac{S_j^{[k]}(t)-S_i^{[k]}(t)}{R_{ij}} + \frac{T_A[k]-S_i^{[k]}(t)}{R_{0i}}\mathbb{1}_A(i)+\frac{u_i[k]}{\tau}\mathbb{1}_H(i)+\frac{o_i[k]}{\tau}\mathbb{1}_L(i),$$

where $\mathbb{1}_A(i)$ is an indicator function that takes value 1 if zone i is connected to the ambient environment and 0 otherwise; $\mathbb{1}_H(i)$ denotes the indicator function that takes value 1 if zone i is directly served by the HVAC system and 0 otherwise; and $\mathbb{1}_L(i)$ is an indicator function that takes value 1 if zone i has the internal loads and 0 otherwise; Assuming zone i is connected to the ambient environment, and dividing both sides by C_i, we have:

$$\begin{aligned}\frac{dS_i^{[k]}(t)}{dt}=&\sum_{j\in\mathcal{N}_i}\frac{S_j^{[k]}(t)-S_i^{[k]}(t)}{C_iR_{ij}}+\frac{T_A[k]-S_i^{[k]}(t)}{C_iR_{0i}}\\ &+\frac{u_i[k]}{C_i\tau}\mathbb{1}_H(i)+\frac{o_i[k]}{C_i\tau}\mathbb{1}_L(i)\\ =&\frac{T_A[k]}{C_iR_{0i}}-S_i^{[k]}(t)\sum_{j\in\mathcal{N}_{i0}}\frac{1}{C_iR_{ij}}+\sum_{j\in\mathcal{N}_i}\frac{S_j^{[k]}(t)}{C_iR_{ij}}\\ &+\frac{u_i[k]}{C_i\tau}\mathbb{1}_H(i)+\frac{o_i[k]}{C_i\tau}\mathbb{1}_L(i),\end{aligned}$$

which further implies that:

$$\begin{aligned}&\frac{dS_i^{[k]}(t)}{dt}+S_i^{[k]}(t)\sum_{j\in\mathcal{N}_{i0}}\frac{1}{C_iR_{ij}}\\ &=\frac{T_A[k]}{C_iR_{0i}}+\sum_{j\in\mathcal{N}_i}\frac{S_j^{[k]}(t)}{C_iR_{ij}}+\frac{u_i[k]}{C_i\tau}\mathbb{1}_H(i)+\frac{o_i[k]}{C_i\tau\mathbb{1}_L(i)}.\end{aligned}$$

By letting $a_i\triangleq\sum_{j\in\mathcal{N}_{i0}}(C_iR_{ij})^{-1}$, $i=1,\dots,N$, and multiplying both sides by e^{a_it}, we have that:

$$\begin{aligned}&e^{a_it}\frac{dS_i^{[k]}(t)}{dt}+e^{a_it}a_iS_i^{[k]}(t)=e^{a_it}\frac{T_A[k]}{C_iR_{0i}}\\ &+e^{a_it}\sum_{j\in\mathcal{N}_i}\frac{S_j^{[k]}(t)}{C_iR_{ij}}+e^{a_it}\frac{u_i[k]\mathbb{1}_H(i)+o_i[k]\mathbb{1}_L(i)}{C_i\tau}.\end{aligned}\quad(44)$$

Integrating (44) from 0 to τ yields:

$$\begin{aligned}e^{a_it}S_i^{[k]}(t)\Big|_0^\tau=&\frac{T_A[k]}{a_iC_iR_{0i}}\int_0^\tau a_ie^{a_i}dt\\ &+\sum_{j\in\mathcal{N}_i}\frac{1}{a_iC_iR_{ij}}\int_0^\tau a_ie^{a_it}S_j^{[k]}(t)dt\\ &+\frac{u_i[k]\mathbb{1}_H(i)+o_i[k]\mathbb{1}_L(i)}{a_iC_i\tau}\int_0^\tau a_ie^{a_it}dt.\end{aligned}$$

Noting that $S_i^{[k]}(0)=T_i[k]$ and $S_i^{[k]}(\tau)=T_i[k+1]$, we have:

$$\begin{aligned}e^{a_i\tau}T_i[k+1]-T_i[k]=&\frac{T_A[k]}{a_iC_iR_{0i}}(e^{a_i\tau}-1)+\\ &\underbrace{\sum_{j\in\mathcal{N}_i}\frac{1}{a_iC_iR_{ij}}\int_0^\tau a_ie^{a_it}S_j^{[k]}(t)dt}_{\langle a\rangle}+\\ &\underbrace{\frac{u_i[k]\mathbb{1}_H(i)+o_i[k]\mathbb{1}_L(i)}{a_iC_i\tau}\int_0^\tau a_ie^{a_it}dt}_{\langle b\rangle}.\end{aligned}\quad(45)$$

Now, consider the two terms $\langle a\rangle$ and $\langle b\rangle$ in (45). For $\langle a\rangle$, we have:

$$\begin{aligned}\langle a\rangle=&\frac{1}{a_iC_iR_{ij}}\int_0^\tau a_ie^{a_it}S_j^{[k]}(t)dt\\ \overset{(c)}{=}&\frac{1}{a_iC_iR_{ij}}\left[e^{a_it}S_j^{[k]}(t)\Big|_0^\tau-\int_0^\tau e^{a_it}\frac{T_j[k+1]-T_j[k]}{\tau}dt\right]\\ \overset{(d)}{=}&\frac{1}{a_iC_iR_{ij}}\Bigg[e^{a_i\tau}T_j[k+1]-T_j[k]\\ &-\frac{T_j[k+1]-T_j[k]}{\tau}\int_0^\tau e^{a_it}dt\Bigg]\\ =&\frac{1}{a_iC_iR_{ij}}\Bigg[e^{a_i\tau}T_j[k+1]-T_j[k]\\ &-\frac{T_j[k+1]-T_j[k]}{a_i\tau}(e^{a_i\tau}-1)\Bigg],\end{aligned}\quad(46)$$

where (c) follows from integration by parts and the fact that $S_j^{[k]}(t)$ being a linear function passing through $T_j[k]$ and $T_j[k+1]$; and (d) follows from that $\frac{T_j[k+1]-T_j[k]}{\tau}$ is a constant independent of t. On the other hand, for $\langle b\rangle$, it is easy to see that

$$\begin{aligned}\langle b\rangle=&\frac{u_i[k]\mathbb{1}_H(i)+o_i[k]\mathbb{1}_L(i)}{a_iC_i\tau}\int_0^\tau a_ie^{a_it}dt\\ =&\frac{u_i[k]\mathbb{1}_H(i)+o_i[k]\mathbb{1}_L(i)}{a_iC_i\tau}(e^{a_i\tau}-1).\end{aligned}\quad(47)$$

Substituting (46) and (47) into (45), we have:

$$\begin{aligned}&e^{a_i\tau}T_i[k+1]-T_i[k]=\frac{T_A[k]}{a_iC_iR_{0i}}(e^{a_i\tau}-1)\\ &+\frac{u_i[k]\mathbb{1}_H(i)+o_i[k]\mathbb{1}_L(i)}{a_iC_i\tau}(e^{a_i\tau}-1)\\ &+\sum_{j\in\mathcal{N}_i}\frac{\big((a_i\tau-1)e^{a_i\tau}+1\big)T_j[k+1]-(a_i\tau-e^{a_i\tau}+1)T_j[k]}{a_i^2C_iR_{ij}\tau}.\end{aligned}\quad(48)$$

Rearranging and collecting terms in (48), we have

$$\begin{aligned}&e^{a_i\tau}T_i[k+1]-\sum_{j\in\mathcal{N}_i}\left[\frac{(a_i\tau-1)e^{a_i\tau}+1}{a_i^2C_iR_{ij}\tau}\right]T_j[k+1]\\ &-T_i[k]+\sum_{j\in\mathcal{N}_i}\left[\frac{a_i\tau-e^{a_i\tau}+1}{a_i^2C_iR_{ij}\tau}\right]T_j[k]\\ &-\frac{e^{a_i\tau}-1}{a_i\tau}(u_i[k]\mathbb{1}_H(i)+o_i[k]\mathbb{1}_L(i))=T_A[k]\frac{e^{a_i\tau}-1}{a_iC_iR_{0i}}\end{aligned}$$

Lastly, by letting coefficients p_i, p_{ij}, q_{ij}, r_i, and s_i be defined as in (11)–(15), it then follows that:

$$p_i T_i[k+1] - \sum_{j \in \mathcal{N}_i} p_{ij} T_j[k+1] - T_i[k] + \sum_{j \in \mathcal{N}_i} q_{ij} T_j[k] \\ - r_i (u_i[k] \mathbb{1}_H(i) + o_i[k] \mathbb{1}_L(i)) = s_i T_A[k],$$

which is the same as stated in (16). This completes the proof.