A Switched Dynamic Programming Approach towards Optimal Control of Multiple Rooftop Units

Vamsi Putta[1], Donghun Kim[2], Jie Cai[2], Jianghai Hu[1], James E. Braun[2]

Abstract— **Buildings served by multiple rooftop units (RTUs) are widely prevalent in retail and food industries. Till date, most of such buildings utilize plain feedback or heuristic strategies to determine the RTU operations. We present a switched dynamic programming based approach for scheduling rooftop unit operations in an optimal manner. Using a mid-size restaurant served by multiple RTUs as an example, a switched affine dynamic system model is first presented. The RTU coordination control problem is then formulated as a switched optimal control problem with nontrivial switching cost. A dynamical programming based method is proposed for the solution of the switched optimal control problem, which has reasonable computational complexity and can be implemented either as a one-shot solution or as part of a model predictive control algorithm. The proposed method is evaluated through simulations, and its strengths and limitations are discussed.**

I. Introduction

Rooftop Units (RTUs) are packaged air handling units that regulate the temperature and circulate air as part of a Heating, Ventilation and Air-Conditioning system (HVAC). Rooftop Units are among the most prevalent HVAC systems in commercial buildings in the United States. Most commercial buildings in the retail and food-processing industries are served by several individually controlled RTUs. In a majority of such buildings, each RTU's operation is controlled via a temperature setpoint manually adjustable at a corresponding thermostat location; and the RTUs are cycled on and off to attain the temperature setpoint in different zones of the building. Heuristic control strategies of individual RTUs without considering their coordination could lead to wasteful scenarios during the operating hours where different RTUs compete (by simultaneous heating and cooling) to maintain their respective temperature setpoints, resulting in higher energy expenditure.

A building with multiple RTU units operating in on/off fashion can be modeled as a switched affine system (SAS). SASs are a class of hybrid dynamical systems comprising of multiple affine dynamical subsystems whose control is specified by a (discrete) switching sequence in addition to a continuous control input. A building with multiple RTUs operates in a number of distinct modes, one corresponding to an ON/OFF configuration of the RTUs. In each mode, the building thermal dynamics can be modeled by an affine dynamics obtained from, e.g., a RC network approximation. The mode sequence over any given time horizon specifies the operating schedule of the RTUs. An optimal control problem can then be formulated to find the operation schedule (and the continuous control, if applicable) that minimizes the energy cost of the RTUs while maintaining comfort for building occupants. Another cost that needs to be considered is the loss in lifetime of the RTUs due to their frequent turning on and off, which can be modeled as switching cost in the SAS optimal control problem formulation.

Dynamic programming has been one of the primary methods for solving optimal control problems, including those for switched and hybrid systems [1]. For a switched optimal control problem such as the RTU coordination problem under consideration, a key challenge lies in its combinatorial nature: the number of operation schedules that need to be considered increases exponentially with the prediction time horizon. Thus, straightforward application of dynamic programming results in exponential complexity growth, making the problem intractable for long time horizons [2]. Mitigations in the form of relaxation [1] and tree pruning [3] have been proposed to reduce the growth of complexity. An application of relaxed dynamic programming with stability analysis to the Model-based Predictive Control (MPC) of switched linear systems was presented in [4]. Nevertheless, practical applications of these methods remain few to date.

In this paper, a model-based predictive control algorithm is proposed for the coordination of multiple RTUs with varying efficiency in a multi-zone building. MPC algorithms have increasingly become attractive options in building control [5], [6] due to their ability to utilize real-time weather and occupant information to minimize energy consumption. In the proposed solution, the RTU coordination problem over a given time horizon is formulated as a discrete-time switched affine quadratic regulation (SAQR) problem with mode-dependent switching costs and solved by the dynamic programming method with complexity reduction techniques. Simulation results show that the proposed approach can lead to reduced energy expenditure of RTUs through a better coordination among them.

This paper is organized as follows. In Section II, we describe a building that will serve as a motivating example of this study. The general model framework and optimal control problem formulation are presented in Section III and Section IV, respectively. In Section V, the proposed dynamic programming solution method is summarized. Controller implementation using dynamic programming methods are discussed in Section VI. Simulation results are presented and discussed in Section VII. Finally, Section VIII contains some concluding remarks.

II. Case Study Building

In this section, we describe a building that will serve as the motivating example and the testing case for the proposed method. The building under consideration, the Harvest Grill Restaurant, is a medium-size restaurant in the Philadelphia area that is served by 4 RTUs of varying capacity and efficiency. Fig. 1 depicts the internal layout of the restaurant, with the numbers indicating the locations of the thermostats for the RTUs. RTU 1 feeding the main dining area is the largest and the most efficient unit while RTU 2,3 and 4 are identical smaller units serving the smaller zones. The capacity and efficiency (COP) of these RTUs are summarized in Table I.

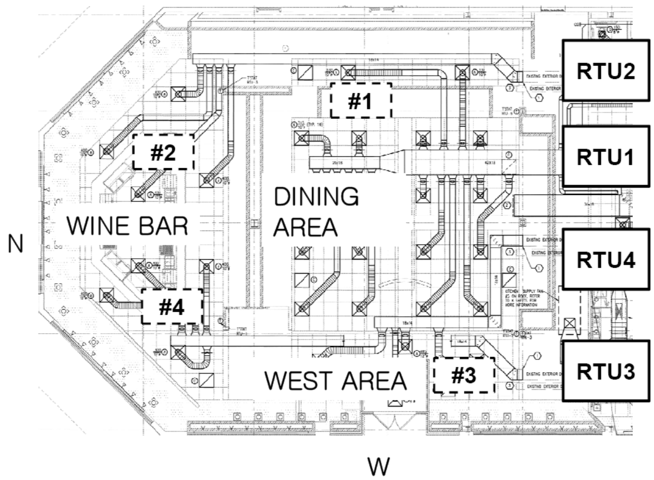


Fig. 1: Layout of Harvest Grill Restaurant.

RTU	Rated Cooling	Energy consumption	COP
RTU 1	52.74 kW	15.06 kW	3.5
RTU 2	14.60 kW	5.65 kW	2.58
RTU 3	14.60 kW	5.65 kW	2.58
RTU 4	14.60 kW	5.65 kW	2.58

TABLE I: RTU Specifications of Harvest Grill Restaurant.

In the existing configuration of the restaurant on site, air temperature is sensed at each of the four thermostat locations, and the corresponding RTU cycles its compressor ON or OFF based on the deviation of the measured temperature from a user specified setpoint without coordinating with other RTUs. All the RTU fans are constantly operational; thus each RTU can be controlled only via its ON/OFF status. In this study, it is assumed that each RTU provides cool air at a given supply air temperature when they are ON and air at ambient temperature when OFF[1]. This assumption can simplify significantly the effect of the RTU operation sequence on the building thermal dynamics.

A high-fidelity Computational Fluid Dynamics (CFD) coupled model of the restaurant's building envelope was developed and validated based on the approach presented in [7]. However, the resulting model had a high state dimension for controller design purposes. Based on this model, a four-dimensional linear system model (inverse model) was developed using the subspace identification (n4sid) method, where the training data was obtained by simulating a typical control strategy for a one-week period designed to keep the zone temperatures (as sensed by the thermostats) at a comfortable level between 21^oC and 26^oC. Such input data is representative of the data available on-site. The TMY3 weather profile for June 2014 was used to model the ambient conditions, while a typical occupancy profile was assumed in the training process. The resulting inverse model was validated by using one week of the July 2014 weather profile. When compared to the CFD coupled model, the trained linear model had an RMSE of 3^oC across all zones over the week. When restricted to the occupancy period the RMSE decreased to 1.1^oC. Figure 2 depicts the dining area temperature measurement during the validation period. In the next section, we describe how the obtained inverse model can be formulated as a switched affine system.

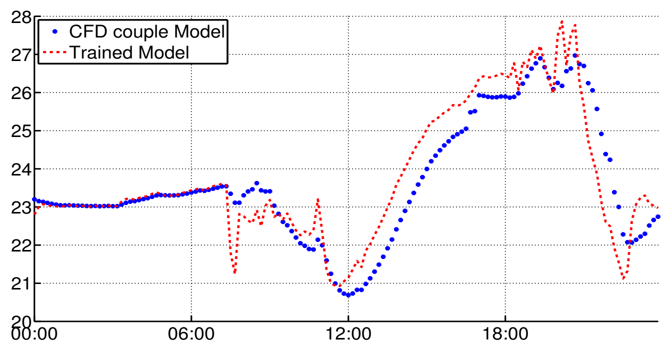


Fig. 2: Temperature trajectory during model validation

III. Switched Affine System Model

We begin by describing a general state space model of the building envelope. Distinguishing between controllable inputs and exogenous inputs, the building thermal dynamics can be modeled by

$$x(t+1) = Ax(t) + Bu_{\sigma(t)} + Fw(t), \qquad (1)$$
$$y(t) = Cx(t), \quad t = 0, 1, \dots .$$

Here, the state $x(t) \in \mathbb{R}^n$ is a vector consisting of the temperatures of the relevant wall and air nodes in the CFD model. In the inverse model, $x(t)$ is a lower-dimensional linear transform of such temperatures. The operation schedule $\{\sigma(t)\}_{t=0,1,\dots}$ is sequence of variables with a finite set Σ of possible values, one corresponding to each possible combined ON/OFF status of all the RTUs. For the case study building with four RTUs, Σ consists of a total of 16 possible values (or *modes*). The uncontrollable input $w(t) \in \mathbb{R}^{n_w}$ models external heat gains from solar radiation, as well as internal gains from lighting and occupants. The controllable input $u_{\sigma(t)} \in \mathbb{R}^{n_u}$ is a vector of zeros and ones with each element corresponding to the ON/OFF status of an RTU for the mode $\sigma(t) \in \Sigma$ at time t. The output variable $y(t) \in \mathbb{R}^p$ contains the air temperature measurements at the thermostat locations. A, B, F, C are constant matrices of proper dimensions.

[1]Strictly speaking, this assumption is not true due to the dependence of the supply air temperature on the ambient condition. However, for the relatively small simulation time horizon in this study, the errors incurred will be small.

The above model can be simplified to

$$x(t+1) = Ax(t) + b_{\sigma(t)} + Fw(t), \qquad (2)$$
$$y(t) = Cx(t), \quad t = 0, 1, \ldots.$$

Here, $b_{\sigma(t)} = Bu_{\sigma(t)} \in \mathbb{R}^n$ denotes the contribution of all the RTUs to the envelope dynamics, whose values depend on the current RTU ON/OFF configuration $\sigma(t)$. Due to the presence of the $Fw(t)$ term and the mode sequence $\{\sigma(t)\}$ as the control input, this is an instance of the switched affine systems (SASs).

To facilitate the study of the ensuing temperature tracking problem with possibly time-varying setpoint offsets, the above SAS model can be reduced to a simpler Switched Linear System (SLS) model through a standard homogenization procedure, if the exogenous input $w(t)$ is assumed to be known or predictable within the given time horizon. To this end, assume that $x_{set}(t)$ is a setpoint (reference) trajectory of the state and $y_{set}(t) = Cx_{set}(t)$ the corresponding setpoint output trajectory. Define the augmented and offset state and the offset output as

$$\hat{x}(t) := \begin{bmatrix} x(t) - x_{set}(t) \\ 1 \end{bmatrix}, \quad \hat{y}(t) := y(t) - y_{set}(t), \qquad (3)$$

respectively. Then (2) can be written as

$$\hat{x}(t+1) = \hat{A}_\sigma(t)\hat{x}(t), \qquad (4)$$
$$\hat{y}(t) = \hat{C}\hat{x}(t),$$

where $\hat{A}_\sigma(t) \in \mathbb{R}^{(n+1)\times(n+1)}$ and $\hat{C} \in \mathbb{R}^{p\times(n+1)}$ are

$$\hat{A}_\sigma(t) = \begin{bmatrix} A & b_{\sigma(t)} + Fw(t) + Ax_{set}(t) - x_{set}(t+1) \\ 0 & 1 \end{bmatrix}$$
$$\hat{C} = \begin{bmatrix} C & 0 \end{bmatrix}.$$

Note that the system in (4) is now a homogeneous switched linear system. By assumptions, $\hat{A}_\sigma(t)$ is known or predictable with values dependent on the mode sequence $\sigma(t)$ within the given time horizon.

IV. Problem Formulation

For the building system described in the previous section, the optimal RTU coordination problem to be formulated in this section is an optimal control problem of the SAS (2) or the SLS (4) whose objective is to minimize the operational cost while maintaining occupant comfort during a given time horizon. For simplicity, we assume in this section that the current time is $t = 0$ and the given time horizon is $t = 0, 1, \ldots, k$. We first discuss in the following three factors that contribute to the cost function over this time period.

a) Energy Cost.: For a building with r RTUs, each mode $\sigma \in \Sigma$ can be represented by a binary string $\sigma_1\sigma_2\cdots\sigma_r$ where $\sigma_j = 0$ and 1 represents the OFF and ON status of RTU $j \in \{1, \ldots, r\}$, respectively. Assume that the power consumed by RTU j is p_j when it is ON and 0 when it is off. Then the total energy consumed by the RTUs during the period $t = 0, 1, \ldots, k$ is given by

$$J_e(k) = \sum_{t=0}^{k} \lambda(t) p_{\sigma(t)} \Delta,$$

where $\lambda(t)$ is the (predicted) utility price at time t, Δ is the sampling time, and $p_{\sigma(t)}$ is the total power consumed by the RTUs in mode $\sigma(t)$:

$$p_{\sigma(t)} = \sum_{j=1}^{r} \sigma_j(t) p_j.$$

b) Comfort Penalty.: Another factor the needs to be considered in the cost function is the occupant comfort. There are various existing metrics to measure occupant (dis)comfort, e.g., Percentage of People Dissatisfied (PPD) and Predicted Mean Vote (PMV) [8], [9]. In this paper, a simplified approach is adopted[2]: we assume that a reference trajectory of the output (i.e., setpoint temperatures at the thermostat locations) $y_{set}(t) = Cx_{set}(t)$ is given over the time horizon $t \in \{0, \ldots, k\}$; and deviation from it at any time t will incur a penalty $[y(t) - y_{set}(t)]^T H(t)[y(t) - y_{set}(t)]$ for some positive semidefinite matrix $H(t)$. Thus, the total comfort penalty cost over the time period is

$$J_c(k) = \sum_{t=0}^{k} [y(t) - y_{set}(t)]^T H(t)[y(t) - y_{set}(t)].$$

c) Switching Cost.: In practice, an RTU has a finite lifespan, which may be shortened by frequently turning it on and off. In addition, turning several RTUs on at the same time may lead to large spikes in power demand which could result in excessive demand charges in the utility bill. Thus, the optimal control problem formulation needs to take into account the cost associated with the RTUs switching status at each time step. Specifically, for each RTU $j \in \{1, \ldots, r\}$, denote by c_{01}^j (resp. c_{10}^j) the cost associated with it switching from OFF to ON (resp. from ON to OFF) at each time step. Let $\sigma(t)$ and $\sigma(t+1)$ be two consecutive modes. Then the cost of switching from $\sigma(t)$ to $\sigma(t+1)$ is

$$c_{\sigma(t),\sigma(t+1)} := \sum_{j=1}^{r} c^j_{\sigma_j(t),\sigma_j(t+1)}.$$

More general switching costs can be introduced, e.g., to penalize multiple RTUs turning on at the same time, which are omitted here. Thus, the total switching cost during the time period $t \in \{0, \ldots, k\}$ is

$$J_s(k) = \sum_{t=0}^{k} c_{\sigma(t),\sigma(t+1)}$$

where for notational simplicity later on we have assumed that $\sigma(k+1) = \sigma(k)$, i.e., there is no switching after the last time step.

Remark 1: Another formulation of the switching cost is $\tilde{J}_s(k) = \sum_{t=0}^{k} c_{\sigma(t-1),\sigma(t)}$, which includes the switching cost from $\sigma(-1)$ to $\sigma(0)$. Using this formulation, the cost will differ only slightly from $J_s(k)$ when time horizon is long; however, the value function of the RTU coordination problem to be defined later on will have a much higher

[2]In our previous work [10] we have shown that the simplified metric gives a good approximation of the PPD and PMV within a certain range of the thermal conditions

complexity as it depends on both current state and previous mode.

d) Optimal Control Problem Formulation.: With the costs defined above, the optimal RTU coordination problem can be formulated as the following optimal control problem:

$$\begin{aligned} &\text{minimize} \;\; J(k) = J_e(k) + J_c(k) + J_s(k) \\ &\text{subject to} \;\; \text{system dynamics (2) for } t = 0, 1, \dots, k. \end{aligned}$$

Note that the optimal control to be solved is the RTU operation schedule $\sigma(t)$ over the horizon $t \in \{0, \dots, k\}$ with the understanding that $\sigma(k+1) = \sigma(k)$.

By using the augmented and offset state and output defined in (3), the SAS dynamics (2) is simplified to the SLS dynamics (4); the cost function $J(k)$ is also reduced to a quadratic one,

$$J(k) = \sum_{t=0}^{k} \hat{x}(t)^T \widehat{H}(t) \hat{x}(t),$$

where $\widehat{H}(t) \in \mathbb{R}^{(n+1)\times(n+1)}$ is defined as

$$\widehat{H}(t) = \begin{bmatrix} C^T H(t) C & 0 \\ 0 & \lambda(t) p_{\sigma(t)} \Delta + c_{\sigma(t),\sigma(t+1)} \end{bmatrix}.$$

As a result, the optimal RTU coordination problem can be equivalently formulated as

$$\begin{aligned} \text{minimize} \;\; & J(k) = \sum_{t=0}^{k} \hat{x}(t)^T \widehat{H}(t) \hat{x}(t) && (5) \\ \text{subject to} \;\; & \hat{x}(t+1) = \hat{A}_\sigma(t) \hat{x}(t), \\ & \hat{y}(t) = \hat{C} \hat{x}(t), \quad t = 0, 1, \dots, k. \end{aligned}$$

Note that problem (5) above can be considered as a generalized version of the switched linear quadratic regulation (SLQR) problem studied in [2] in that the quadratic matrix $\widehat{H}(t)$ for the cost at time t depends not only on the current mode $\sigma(t)$ but also on the next mode $\sigma(t+1)$.

V. Dynamic Programming Solution

The optimal control problem (5) formulated in the previous section can be solved using a dynamic programming method to be presented in this section, which is a (slight) generalization of the algorithm proposed in [3]. Denote by $V_s(\hat{x})$ the value function (cost-to-go) of the problem (5) over the time horizon $\{s, s+1, \dots, k\}$:

$$V_s(\hat{x}) := \min_{\sigma(t),\, s \le t \le k} \left\{ \sum_{t=s}^{k} \hat{x}(t)^T \widehat{H}(t) \hat{x}(t) \,\middle|\, \hat{x}(s) = \hat{x} \right\}$$

for $\hat{x} \in \mathbb{R}^{n+1}$ and $s \in \{0, 1, \dots, k+1\}$. Then $V_s(\hat{x})$ satisfies the Bellman equation

$$V_s(\hat{x}) = \min_{\sigma(s) \in \Sigma} \left\{ \hat{x}^T \widehat{H}(s) \hat{x} + V_{s+1}\left(\hat{A}_\sigma(s) \hat{x} \right) \right\} \quad (6)$$

for all $\hat{x} \in \mathbb{R}^{n+1}$ and $s \in \{0, 1, \dots, k\}$, with zero terminal cost $V_{k+1}(\cdot) \equiv 0$ due to our assumption that $\sigma(k+1) = \sigma(k)$. Note that, in (6), both terms in the bracket depend on the decision variable $\sigma(s)$ as both matrices $\widehat{H}(s)$ and $\hat{A}_\sigma(s)$ are $\sigma(s)$-dependent.

It should be pointed out that, in the Bellman equation (6), the matrix $\widehat{H}(s)$ also depends on $\sigma^*(s+1)$, which is the optimal mode when starting from the state $\hat{x}(s+1) := \hat{A}_\sigma(s) \hat{x}$ at the next time step $s+1$:

$$\begin{aligned} \sigma^*(s+1) = \arg\min_{\sigma(s+1) \in \Sigma} \Big\{ & \hat{x}(s+1)^T \widehat{H}(s+1) \hat{x}(s+1) + \\ & V_{s+2}\left(\hat{A}_\sigma(s+1) \hat{x}(s+1) \right) \Big\}, \quad (7) \end{aligned}$$

for $s \in \{0, 1, \dots, k-1\}$. Note that $\sigma^*(s+1)$ itself depends on $\sigma(s)$. When $s = k$, $\sigma^*(k+1) = \sigma(k)$.

To sum up, the special structure of the $\widehat{H}(t)$ matrix in the optimal RTU coordination problem renders the iteration (6) a less straightforward process than the conventional Bellman iterations in that the control decision $\sigma(s)$ affects the running cost $\hat{x}^T \widehat{H}(t) \hat{x}$ in a complicated way through the dependency of $\widehat{H}(t)$ on $\sigma^*(s+1)$ and hence ultimately on $\hat{A}_\sigma(s)\hat{x}$.

Carrying out the iteration (6) for $s = k, k-1, \dots, 0$ with the terminal conditions $V_{k+1}(\cdot) \equiv 0$ and $\sigma^*(k+1) = \sigma(k)$, we can obtain all the value functions $V_s(\cdot)$, in particular, $V_0(\cdot)$. The optimal cost for problem (5) is then $V_0(\hat{x}_0)$ where $\hat{x}_0 = \begin{bmatrix} x(0)^T & 1 \end{bmatrix}^T$. The optimal mode sequence $\{\sigma^*(t)\}_{t=0,\dots,k}$ resulting in the optimal cost can be recovered by a forward iteration, yielding the optimal operation schedule for all the RTUs.

A. Representation of Value Functions

For SLQR problems without switching cost, it was shown in [2] that their value functions are the pointwise minimum of finite families of quadratic functions, hence piecewise quadratic in themselves. For the RTU coordination problem (5) with switching cost, such a property still holds.

Proposition 1: For each $s \in \{0, 1, \dots, k\}$, the value function $V_s(\cdot)$ can be represented as

$$V_s(\hat{x}) = \min_{\hat{Q} \in \mathcal{Q}_s} \hat{x}^T \hat{Q} \hat{x} \quad (8)$$

for some finite set $\mathcal{Q}_s$ of positive semidefinite matrices in $\mathbb{R}^{(n+1)\times(n+1)}$.

Proof: Only a sketch of the proof is given below. For a fixed operation schedule, the dynamics (4) becomes a linear time-varying system, and all the quadratic matrices $\widehat{H}(t)$ are fixed, resulting in quadratic costs-to-go. For the switched optimal control problem (5), each value function is the minimal of such costs-to-go over all possible operation schedules. This proves the desired result. ■

Using the representation (8), the value function iterations (6) then reduce to iterative procedures for obtaining the sets $\mathcal{Q}_k$, $\mathcal{Q}_{k-1}$, …, $\mathcal{Q}_0$ as follows

$$\mathcal{Q}_s = \rho_\Sigma^s(\mathcal{Q}_{s+1}), \quad s = k, k-1, \dots, 0.$$

with the terminal condition $\mathcal{Q}_{k+1} = \{0\}$. Here, ρ_Σ^s is the *switched Riccati mapping* at time s defined by

$$\rho_\Sigma^s(\mathcal{Q}_{s+1}) := \left\{ \rho_{\sigma,\sigma'}^s(Q) \,\middle|\, \sigma, \sigma' \in \Sigma,\, Q \in \mathcal{Q}_{s+1} \right\}, \quad (9)$$

while $\rho^s_{\sigma,\sigma'}$ for $\sigma, \sigma' \in \Sigma$ is the Riccati mapping so that

$$\rho^s_{\sigma,\sigma'}(Q) = \hat{A}_\sigma(s)^T Q \hat{A}_\sigma(s) + \widehat{H}(s) \in \mathbb{R}^{(n+1)\times(n+1)} \quad (10)$$

for any $Q \in \mathbb{R}^{(n+1)\times(n+1)}$. Note that in (10), $\hat{A}_\sigma(s)$ depends on σ while $\widehat{H}(s)$ depends on the mode pair (σ, σ').

Remark 2: In the case there is no switching cost, $\widehat{H}(s)$ depends only on the current mode $\sigma(s)$; hence σ' can be dropped from both of the definitions (9) and (10).

The switched Riccati mapping (9) provides a way of computing the sets $\mathcal{Q}_s$, hence the value functions $V_s(\cdot)$, iteratively. However, the complexity of the representation (8) as measured by the cardinality of the set $\mathcal{Q}_s$ grows exponentially with the iterations. Several complexity reduction techniques were proposed in [3], including pruning at each step redundant matrices that do not contribute to the minimum in (8), and relaxed dynamical programming methods that remove almost redundant matrices at the expense of accuracy. These techniques after slight modification can be applied to reduce the computational complexity of the iteration (9). For more details, the interested readers can refer to [3].

VI. Control Implementations

The dynamic programming algorithm proposed in Section V can be implemented in several different ways, depending on the (conflicting) demands of real-time computation and solution sub-optimality.

A. MPC Implementation

Model Predictive Control incorporates estimates of the exogenous inputs to solve the optimal control problem (5) at every time instant and implements only the first step of the obtained optimal control. By updating the state and exogenous input estimates at each time step, MPC implementation yields solutions that can adapt quickly to changes in inside/outside conditions and are more robust to errors in prediction of exogenous inputs, a desirable feature in building applications as these inputs are usually stochastic in nature. Another benefit is that the time horizon of the MPC algorithm can be chosen properly to achieve a balance between computation complexity and solution sub-optimality: a shorter time horizon can significantly reduce the complexity of the dynamic programming algorithms proposed in Section V, though at the expense of yielding relatively short-sighted solutions with subpar performance in the long term; and vice versa.

B. One-Shot Implementation

If accurate prediction of the exogenous inputs $w(k)$ is available over a long time horizon, the optimal control problem (5) can be solved only once, with the resulting optimal RTU operating schedule implemented during the whole lookahead horizon (rather than just its first step as in the MPC case). This approach will generally lead to a better performance compared to moving-horizon type controllers with a shorter prediction horizon, though at the expense of a longer computation time. On the other hand, as the problem becomes a finite-horizon switched discrete-time LQR problem that needs to be solved only infrequently, its solution can be carried out offline at non-critical times such as at night when the optimal control is straightforward, leading to real-time implementation.

C. Quantized Control

At the other extreme, a complete lack of forecast makes it feasible to use only the current measurements of the exogenous inputs. In such cases, the cost function can only reflect the temperature regulation and power consumption over a very short time horizon. Due to the emphasis on the temperature regulation, the problem becomes a discrete optimization problem of load matching, which can be formulated and solved as an instance of the optimal quantized control problem [11]. Generally speaking, the resulting RTU operations will likely be suboptimal over a long period of time.

VII. Simulation Results

In this section, the performance of the proposed algorithms will be tested on the Harvest Grill Restaurant model. The savings in power cost and temperature regulation behaviors for the different implementations of the proposed algorithm will be compared.

The three control implementations discussed in Section VI were simulated using the 4-dimensional black-box model of the Harvest Grill Restaurant described in Section II. The inverse model was used for both control design and plant emulation. A look ahead horizon of 6 hours ($k = 36$ sampling instants with a sampling time of 10 minutes) was used with perfect prediction of the exogenous inputs. The comfort penality quadratic matrix $H(\cdot)$ was chosen to be $7I_4$, where I_4 denotes the 4-dimensional identity matrix. The temperature setpoint y_{set} was chosen to be 23^oC for all zones. Each RTU was assigned a switching cost corresponding to 15% of its power consumption. This reflects a higher penalty in cycling the largest RTU on and off, thus implicitly modeling a demand charge based scenario where consumers are billed on the maximum power usage.

Simulations were performed to obtain the power consumptions of the three control implementations during a one-week period. Performance in maintaining comfort is quantified in terms of the RMSE deviation of the zone temperatures from the setpoints during the occupancy hours, averaged across the four zones. A simple feedback based algorithm, where the RTUs are cycled based on the temperature measurement at the corresponding thermostats was also simulated to serve as the baseline. A deadband of $\pm 0.5^oC$ was included in the control logic to prevent chattering. Tree pruning with a relaxation parameter of $1e^{-2}$ was used to keep the computation feasible in real time for the dynamic programming algorithm. All controllers were designed to operate from $6:00$ am till midnight with the temperature being allowed to float during the remaining times.

Figure 3 and 4 show the performance of the controllers in terms of maintaining comfort and the corresponding RTU operations. As can be seen the controllers are able to

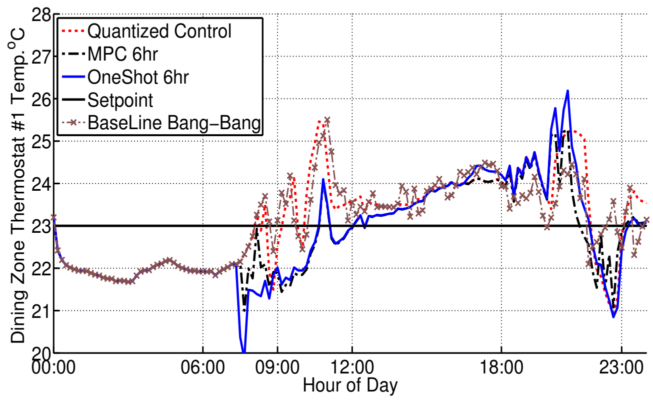


Fig. 3: Temperature variations under different control implementations.

maintaining thermal comfort to an extent. The one-shot and the MPC controllers offer some precooling in anticipation of the upcoming internal gains at the start of the working hours. The conventional bang-bang controller slightly outperforms the other controllers in temperature regulation by virtue of focusing on temperature tracking alone.

During the occupancy period, the quantized controller is observed to exhibit less cycling compared to the look-ahead based controllers. This is due to the choice of the switching cost. Longer look-ahead controllers amortize the switching costs over a longer horizon by saving power costs. The quantized controller is reactive and the relatively high switching costs force it to avoid frequent cycling. The look ahead based controllers turn on the more efficient RTU#1 earlier and run it longer which is reflected in the power savings. The conventional bang-bang controller exhibits more a longer run times across all RTUs. The included deadband helps minimizing short cycling and can be adjusted as necessary.

Table II summarizes the mean energy consumption per day and the temperature regulation across all the zones. The one-shot controller offers the most savings (8%) compared to the baseline bang-bang control. The 6-hour lookahead MPC comes close to matching the savings (7%). The quantized control offers a savings of (3.45%) while avoiding frequent cycling as well. All the three controllers perform reasonably well in thermal regulation with the average deviation being less than $1.5°C$ though the bang-bang controller provides much tighter thermal regulation.

Controller	**Avg. Energy Usage**	**RMSE Temp. Deviation**
One-Shot 6hr lookahead	286.13 kWh/day	$1.45°C$
MPC 6hr lookahead	289.17 kWh/day	$1.29°C$
Quantized control	300.43 kWh/day	$1.3°C$
Bang-Bang control	311.14 kWh/day	$0.9°C$

TABLE II: Summary of various controllers' performance.

Solution Complexity: Due to the combinatorial nature of the optimization problem, the number of matrices required to represent the value function in (5) grows exponentially with the time horizon. For the current simulation, we mitigate this complexity growth by pruning and relaxing the value function computation [2]. Using a relaxation threshold of $\epsilon = 1e^{-2}$ yields approximately 850 matrices in the representation of the value function after 36 iterations (6 hour lookahead with a sampling time of 10 minutes) compared to the theoretical upperbound of 16^{36} matrices. On a 2.4Ghz Intel Core i3 processor based PC, the maximum computation time to solve the MPC problem at any sampling instant was 8.8 minutes which points to real time feasibility considering the 10 minute sampling interval.

VIII. Conclusions

Applying techniques from optimization of switched linear systems to building control enables us to formulate and solve the problem of optimal RTU coordination over sufficiently long lookahead horizons with nontrivial switching cost. Simulation results indicate energy savings while being real time feasible.

One limitation of the proposed strategy is the total cost combines several costs related whose weights can significantly affect the optimal solution. The choices of their weight, currently by trial and error, need to be studied formally. In addition, as with all dynamic programming based algorithms, the proposed algorithm also suffers from the curse of dimensionality. However, this may be alleviated by employing model order reduction techniques and efficient LMI based solvers.

References

[1] A. Rantzer, "Relaxed dynamic programming in switching systems," *IEE Proceedings-Control Theory and Applications*, vol. 153, no. 5, pp. 567–574, 2006.

[2] W. Zhang, J. Hu, and A. Abate, "On the value functions of the discrete-time switched LQR problem," *Automatic Control, IEEE Transactions on*, vol. 54, no. 11, pp. 2669–2674, Nov 2009.

[3] ——, "Infinite horizon switched LQR problems in discrete time: A suboptimal algorithm with performance analysis," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1815–1821, 2012.

[4] D. Gorges, M. Izák, and S. Liu, "Optimal control and scheduling of switched systems," *Automatic Control, IEEE Transactions on*, vol. 56, no. 1, pp. 135–140, 2011.

[5] Y. Ma, J. Matusko, and F. Borrelli, "Stochastic model predictive control for building hvac systems: Complexity and conservatism," 2013.

[6] A. Parisio, D. Varagnolo, M. Molinari, G. Pattarello, L. Fabietti, and K. H. Johansson, "Implementation of a scenario-based mpc for hvac systems: an experimental case study," in *The 19th World Congress of the International Federation of Automatic Control*, 2014.

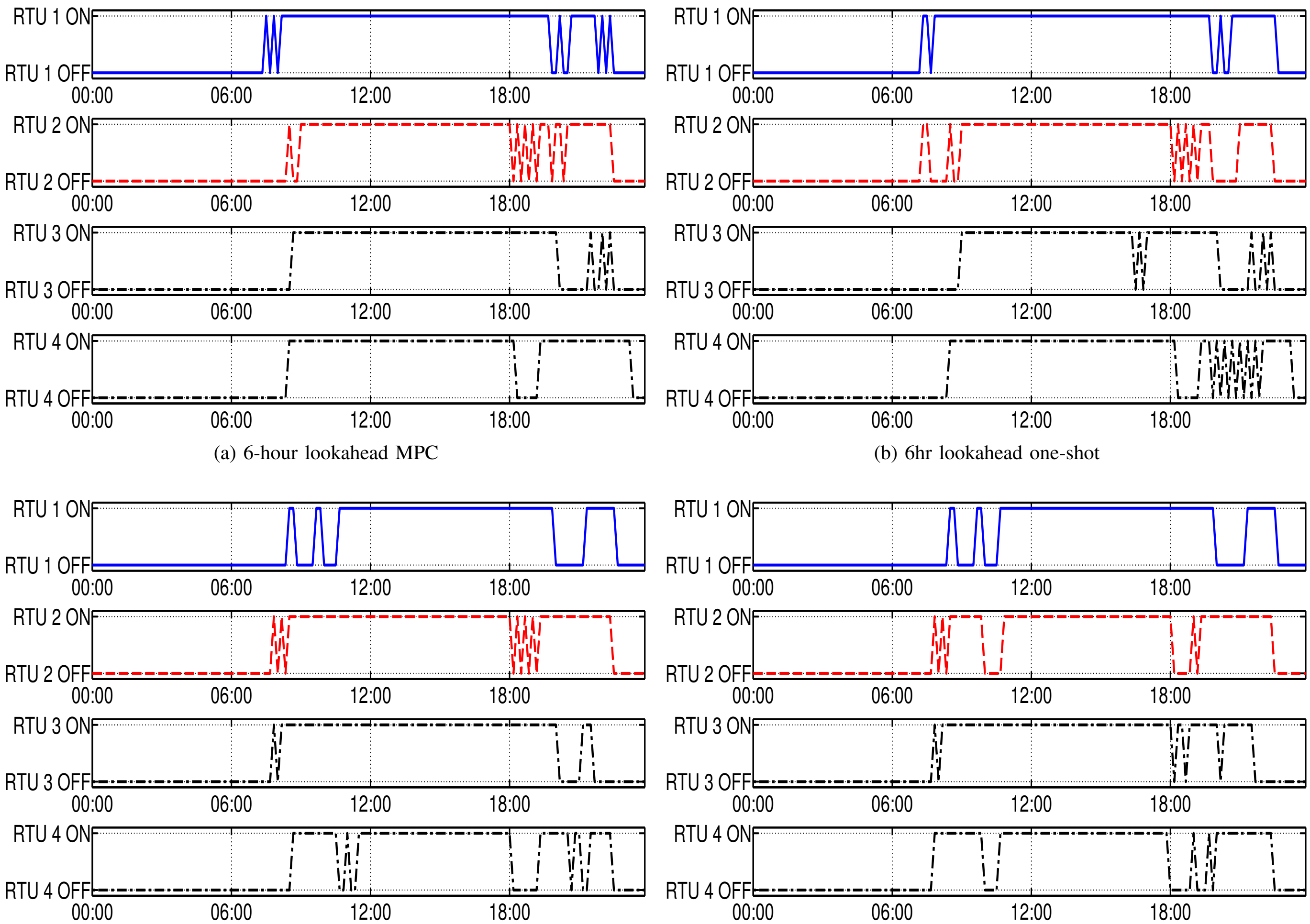
[7] D. Kim, J. Braun, J. Borggaard, E. Cliff, and S. Gugercin, "Coupled cfdbuilding envelope model for the purdue living lab," in *2nd International High Performance Buildings Conference*. Citeseer, 2012.

[8] P. O. Fanger *et al.*, "Thermal comfort. analysis and applications in environmental engineering." *Thermal comfort. Analysis and applications in environmental engineering.*, 1970.

[9] R. De Dear, "Thermal comfort in practice," *Indoor air*, vol. 14, no. s7, pp. 32–39, 2004.

[10] V. Putta, G. Zhu, D. Kim, J. Hu, and J. E. Braun, "Comparative evaluation of model predictive control strategies for a building hvac system," in *Proceedings of the American Control Conference 2013*, 2013, pp. 3461–3467.

[11] S. Pancanti, L. Leonardi, L. Pallottino, and A. Bicchi, "Optimal control of quantized input systems," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, C. Tomlin and M. Greenstreet, Eds. Springer Berlin Heidelberg, 2002, vol. 2289, pp. 351–363.



(a) 6-hour lookahead MPC

(b) 6hr lookahead one-shot

(c) Quantized control with no lookahead

(d) Baseline Bang-Bang control

Fig. 4: Typical RTU operation