Model Predictive Control of Central Chiller Plant With Thermal Energy Storage Via Dynamic Programming and Mixed-Integer Linear Programming

Kun Deng, Member, IEEE, Yu Sun, Sisi Li, Student Member, IEEE, Yan Lu, Member, IEEE, Jack Brouwer, Prashant G. Mehta, Member, IEEE, Meng Chu Zhou, Fellow, IEEE, and Amit Chakraborty, Member, IEEE

Abstract—This work considers the optimal scheduling problem for a campus central plant equipped with a bank of multiple electrical chillers and a thermal energy storage (TES). Typically, the chillers are operated in ON/OFF modes to charge TES and supply chilled water to satisfy the campus cooling demands. A bilinear model is employed to describe the system dynamics of the central plant. A model predictive control (MPC) problem is formulated to obtain optimal set-points to satisfy the campus cooling demands and minimize daily electricity cost. At each time step, the MPC problem is represented as a large-scale mixed-integer nonlinear programming problem. We propose a heuristic algorithm to obtain suboptimal solutions for it via dynamic programming (DP) and mixed integer linear programming (MILP). The system dynamics is linearized along the simulated trajectories of the system. The optimal TES operation profile is obtained by solving a DP problem at every horizon, and the optimal chiller operations are obtained by solving an MILP problem at every time step with a fixed TES operation profile. Simulation results show desired performance and computational tractability of the proposed algorithm.

Note to Practitioners—This work was motivated by the supervisory control need for a campus central plant. Plant operators have to decide a scheduling strategy to mix and match various chillers with a thermal energy storage to satisfy the campus cooling demands, while minimizing the operation cost. This work mathematically characterizes the system dynamics of a campus central plant and establishes a linear model to predict campus cooling load. It proposes a model predictive control (MPC) strategy to optimally schedule the campus central plant based on plant system dynamics and predicted campus cooling load. A heuristic algorithm is proposed to obtain suboptimal solutions for the MPC problem. The effectiveness and efficiency of the proposed approach are well demonstrated for the central plant at the University of California, Irvine.

Index Terms—Central chiller plant, dynamic programming, heuristic algorithm, mixed integer linear programming, model predictive control, supervisory control, thermal energy storage.

I. INTRODUCTION

BUILDINGS are one of the primary consumers of energy in the United States. Buildings are responsible for 41% of primary energy consumption in 2010, compared to 28% by the industrial sector and 31% by the transportation sector [1]. Among all energy consumed by the buildings sector, residential buildings accounted for 54% and commercial buildings accounted for 46%. Among all building energy consumers, heating, ventilation, and air conditioning (HVAC) accounted for the largest share, which represented close to half of all energy consumption [1]. Energy management of HVAC systems becomes essentially important to improve the energy efficiency and reduce the energy cost of buildings.

A central chiller plant is a commonly used HVAC system for a campus with a large number of buildings [2]. The campus cooling loads can be huge and vary in a large range depending on the weather condition and number of buildings. A bank of multiple chillers is usually installed to meet various cooling demands with high operational flexibility. They may have different performance characteristics and are often operated in ON/OFF modes [3], that is, a chiller can be turned on with the maximum mass flow rate of chilled water supply, or be turned off with zero rate. Thermal energy storage (TES) can be used to shape the campus cooling load and reduce the peak one [2]. In the simplest case, TES is a large tank of chilled water. A typical campus central plant with a chiller bank and a TES is depicted in Fig. 1. The chiller bank can either supply chilled water directly to the campus for cooling, or store some of chilled water in TES for later use [4], [5]. Intuitively, TES is charged by the chiller bank at night when electricity rates and cooling demands are both low. During on-peak hours in daytime, it...
A typical control system for the central plant has a two-level control structure: supervisory and local ones [6]–[8]. The goal of supervisory control is to optimize TES and chiller bank operations for minimizing energy consumption or electricity cost while satisfying campus cooling demands. The operation set-points are obtained by taking into account of system level characteristics and interactions among all subsystems. The local control modules individual components, e.g., a relay, valve, or actuator, to track the set-points provided by the supervisory control. A closed-loop feedback control, e.g., PID control, is usually used to achieve the set-point tracking performance and guarantee the robust operation of components. The overall energy efficiency and control effectiveness of a central plant is determined by the performance and coordination of controls in both levels.

Well-designed control systems can help realize the full potential of both energy and cost savings for a central plant [8]–[11]. The focus of this paper is on the design of optimal supervisory control using a model predictive control (MPC) scheme. The main idea of MPC is to use the system model of a central plant to predict its future evolution and obtain its optimal control in a moving horizon basis. The objective is to find optimal set-points by using an MPC scheme to minimize the operational cost while satisfying campus cooling demands and system constraints. At each time step, a model-based optimization problem is solved over a finite horizon to obtain optimal set-points of operating the TES and chiller bank. More details on MPC theory can be found in [12].

For the supervisory control of a central chiller plant, MPC techniques have been investigated since the late 1980s [13], [14]. More early studies on modeling, prediction, and optimization for building HVAC systems can be found in [6], [7], [15], [16] and references therein. With the growing interests in optimal control, many new methods have recently been used to develop MPC schemes for the central chiller plant with TES [17]–[20]. For instance, the work [17] proposed a method by using dynamic programming to solve the optimal chiller sequencing problem. The potential energy savings of the proposed method was demonstrated with three example systems. But the chiller considered in [17] was not equipped with a TES. In [18], [20], the authors considered the MPC problem of a series of chillers equipped with a water tank—a similar setup as this paper. They validated the system model with experimental data and performed closed-loop experimental tests for the proposed control method. To simplify the optimization problem for MPC, they assumed that the mass flow rates of chilled water can be continuously adjusted for chillers, and the charging/discharging profile for TES is fixed or known in advance. Similar assumptions for chillers and TES have also been made in [19].

However, in practice, the operation decisions for chillers are often ON/OFF scheduling sequences and TES operation profile varies depending on the actual cooling demands. The optimization/control frameworks proposed in previous studies do not directly apply for supervisory control of a central chiller plant with TES under practical considerations. To generate optimal ON/OFF decision sequences for chillers and charging/discharging decisions for TES, a mixed-integer nonlinear programming (MINLP) problem can be formulated and solved at each time step using an MPC scheme. The nonlinear constraints result from the system dynamics and integer decision variables are due to the operational constraints for chillers and TES. Since this MINLP problem involves considerable nonlinearity and a large number of integer variables, finding its global optimal solution is nontrivial and computationally expensive [21], [22]. In practice, it is of interest to develop efficient algorithms to approximately solve it.

In [23], we present a heuristic algorithm to search for the suboptimal solutions of the MINLP problem by mixed-integer linear programming (MILP). Its main idea is to explore the special structure of the original problem and solve a set of subproblems with reduced complexity. We formulate an MILP problem to find optimal chiller operations by linearizing the system dynamics with a fixed/given TES operation profile. Compared with [23], we newly propose a dynamic programming (DP)-based algorithm to find optimal TES operations. Now the optimal TES and chiller operations are obtained in two steps. In the first step, we generate an optimal TES operation profile for a fixed horizon via DP with system dynamics approximated by a finite state machine. In the second step, we fix the TES operation profile for a fixed future horizon, and solve for the optimal chiller operations via a reformulated MPC problem. The original MINLP problem is approximated by the reformulated MPC problem with the linearized system dynamics and fixed TES operation profile, which leads to a mixed integer linear program (MILP) at each time step. The nominal trajectory for linearization is obtained by simulating the system with the optimal predicted control trajectory calculated from the last time step MPC problem. The major advantage of the proposed algorithm is that both DP and MILP problems can be solved very efficiently with guaranteed performance. Moreover, the model mismatch due to the linearization can be compensated by the moving horizon scheme used in the MPC scheme. A description of the system model for a chiller bank and TES is provided in Section II. The MPC problem is
formulated in Section III. The DP-based heuristic algorithm for obtaining the optimal TES operation profile is presented in Section IV. The MILP-based heuristic algorithm for obtaining optimal chiller operations is described in Section V. A case study is provided in Section VI, and conclusions are drawn in Section VII.

II. SYSTEM MODELING

The schematic diagram of a typical campus central plant is depicted in Fig. 1, with all pumps being omitted.

A. Nomenclature

We use $dx/dt$ to denote the derivative of variable $x$ with respect to time $t$, and $m$ to denote a rate related input or constant, e.g., chilled water mass flow rate. Time-varying variables are indicated with $(t)$ appended.

- $t$: Time index (s).
- $\rho$: Density of water (kg/m$^3$).
- $c_{pw}$: Specific heat capacity of water (kJ/kg·K).
- $T_{amb}(t)$: Ambient temperature (K).
- $h_{amb}(t)$: Ambient humidity ratio.
- $A$: Cross-layer area of TES tank (m$^2$).
- $i$: Chiller index.
- $n_c$: Total number of chillers.
- $\text{COP}_i$: Coefficient of Performance of the $i$th chiller.
- $\bar{m}_{c,i}$: Chilled water mass flow rate of chiller $i$ (kg/s).
- $T_{c,i}$: Chilled water supply temperature of chiller $i$ (K).
- $T_{cr,i}$: Chilled water return temperature of chiller $i$ (K).
- $W_{c,i}$: Electricity consumed by chiller $i$ (kW).
- $\dot{Q}_{c,i}$: Cooling amount provided by chiller $i$ (kW).
- $\dot{m}_{c,i}$: Chilled water circulation through TES (kg/s).
- $T_a(t)$: Top layer ‘a’ temperature of TES (K).
- $T_b(t)$: Bottom layer ‘b’ temperature of TES (K).
- $\bar{m}_{cb}$: Chilled water mass flow rate of chiller bank (kg/s).
- $T_{cr,b}$: Chilled water return temp. of chiller bank (K).
- $T_{c,b}$: Chilled water supply temp. of chiller bank (K).
- $W_{c,b}$: Electricity consumed by chiller bank (kW).
- $\dot{Q}_{c,b}$: Cooling amount provided by chiller bank (kW).
- $Q(t)$: Cooling load of campus (kW).
- $\dot{m}_{c}$: Chilled water mass flow rate to campus (kg/s).
- $T_s(t)$: Chilled water supply temperature to campus (K).
- $T_r(t)$: Chilled water return temperature from campus (K).

B. Chiller Bank Model

A chiller bank is composed of $n_c$ electrical chillers operated only in ON/OFF modes. When a chiller is ON, it supplies chilled water at a fixed maximum mass flow rate and supply temperature. When OFF, it is shut down and supplies no chilled water. For $i \in \{1, \ldots, n_c\}$, the set-point of the $i$th chiller is a binary decision variable

$$\delta_i = \begin{cases} 1, & \text{if chiller } i \text{ is ON} \\ 0, & \text{if chiller } i \text{ is OFF.} \end{cases}$$

(1)

The amount of cooling provided by chiller $i$ is

$$\dot{Q}_{c,i} = \bar{m}_{c,i} c_{pw} (T_{cr,i} - T_{c,i}).$$

(2)

Note that both $\bar{m}_{c,i}$ and $T_{c,i}$ are fixed constants, and $T_{cr,i}$ varies according to the system dynamics.

The total electricity consumed by a chiller depends on several factors, e.g., supply water mass flow rate $\bar{m}_{c,i}$, generated cooling amount $\dot{Q}_{c,i}$, and circulated water temperature difference $(T_{cr,i} - T_{c,i})$. Thus the chiller electricity consumption mainly depends on temperature difference $(T_{cr,i} - T_{cs,i})$. Here, we simply use a linear regression model to estimate the total electricity consumption of a chiller as well as chilled water pump, condenser water pump, and cooling tower fans. Following [19], the electricity consumption $W_{c,i}$ of the $i$th chiller is modeled as

$$W_{c,i} = a_i (T_{cr,i} - T_{cs,i}) + b_i$$

(3)

where coefficients $a_i$ and $b_i$ can be obtained by regression analysis of the actual measured chiller data. The Coefficient of Performance (COP) is employed as a metric to quantify the energy efficiency of a chiller as follows:

$$\text{COP}_i = \frac{\dot{Q}_{c,i}}{W_{c,i}}.$$  

(4)

The total chilled water mass flow rate supplied by the chiller bank is

$$\dot{m}_{c} = \sum_{i=1}^{n_c} \delta_i \bar{m}_{c,i}.$$  

(5)

The total power consumed by the chiller bank is

$$W_{c} = \sum_{i=1}^{n_c} \delta_i W_{c,i}.$$  

(6)

The total cooling amount provided by the chiller bank is

$$\dot{Q}_{c} = \sum_{i=1}^{n_c} \delta_i \dot{Q}_{c,i}.$$  

(7)

The supply and return temperatures of the chiller bank are denoted by $T_{cs}$ and $T_{cr}$, i.e., for $i \in \{1, \ldots, n_c\}$,

$$\begin{align*}
T_{cs,i} &= T_{cs} \\
T_{cr,i} &= T_{cr}.
\end{align*}$$

(8)
C. TES Model

We employ a stratified two-layer TES model developed in [19], where the state of TES is captured by the top-layer water temperature \( T_w \) and the bottom-layer one \( T_b \). The model is further simplified by ignoring the time delays of the TES tank output. We also assume that the TES tank is always full, implying that the mass flow rate entering the tank is equal to the mass flow rate exiting the tank.

The TES tank operates in two modes. In the charging mode, chilled water from the chiller bank is supplied to both the TES and campus, as shown in Fig. 1, while in the discharging mode chilled water from the TES and chiller is supplied to the campus together. The mode is denoted by a binary variable

\[
\sigma = \begin{cases} 
1, & \text{if } \dot{m}_c > \dot{m}_w \\
0, & \text{if } \dot{m}_c < \dot{m}_w.
\end{cases}
\]  

(9)

When charging, i.e., \( \sigma = 1 \), TES has the following dynamics:

\[
\begin{align*}
T_w &= T_{wc} \\
\dot{m}_w &= \dot{m}_c - \dot{m}_w \\
\dot{m}_w T_{wc} &= \dot{m}_w T_a + \dot{m}_w T_r \\
\rho c_w \frac{dT_w}{dt} &= f_{ac} \dot{m}_w c_{pu}(T_b - T_a) + U_c A(T_b - T_a) \\
\rho c_w \frac{dT_b}{dt} &= f_{bc} \dot{m}_w c_{pu}(T_a - T_b) + U_c A(T_a - T_b).
\end{align*}
\]  

(10)

When discharging, i.e., \( \sigma = 0 \), it has

\[
\begin{align*}
T_r &= T_{cr} \\
\dot{m}_r &= \dot{m}_c - \dot{m}_r \\
\dot{m}_r T_{cr} &= \dot{m}_r T_b + \dot{m}_c T_{cs} \\
\rho c_w \frac{dT_r}{dt} &= f_{ad} \dot{m}_r c_{pu}(T_r - T_a) + U_d A(T_r - T_a) \\
\rho c_w \frac{dT_b}{dt} &= f_{bd} \dot{m}_r c_{pu}(T_a - T_b) + U_d A(T_a - T_b).
\end{align*}
\]  

(11)

Parameters \( f_{ac}, f_{bc}, f_{ad}, \) and \( f_{bd} \) are time constant multipliers, and \( U_c \) and \( U_d \) are inter-layer overall heat transfer coefficients, which can be calibrated using the actually measured TES data.

D. Campus Load Model

The campus cooling load satisfies the following energy balance equation:

\[
\dot{Q} = \dot{m}_w c_{pu}(T_w - T_a).
\]  

(12)

Note that \( \dot{m}_w, T_a, \) and \( T_w \) are taken as system variables, while \( \dot{Q} \) is taken as a given exogenous input to the system from a linear regression model described in the Appendix.

E. Model Summary

By collecting system dynamic equations (1)–(12) and discretizing the system with the sampling time \( \Delta t \), we represent the system dynamics in a state-space form as

\[
\begin{align*}
x(k+1) &= f(x(k), z(k), u(k)) \\
g_1(x(k), z(k), u(k), d(k)) &= 0 \\
g_2(z(k), u(k)) &= 0
\end{align*}
\]  

(13)

where the state variable of the system is denoted by

\[
x = [T_a, T_b]
\]

the intermediate variable is denoted by

\[
z = [W_r, W_{c,i-1}, \ldots, n_r, \dot{Q}_{r,i-1}, \ldots, n_r, \dot{m}_c, T_{cr}, T_s, T_r]
\]

the control variable is denoted by

\[
u = [\sigma, b_1, \ldots, n_r, \dot{m}]
\]

and the input variable is denoted by

\[
d = [\dot{Q}].
\]

Equation (13a) corresponds to the TES dynamic equations (10d), (10e), (11d), and (11e); (13b) corresponds to the TES and campus balance equations (10a)–(10c), (11a)–(11c), and (12); (13c) corresponds to the chiller balance equations (2)–(8).

Besides the system dynamics, we also impose the various constraints on system variables, i.e., there exist bounded sets \( X, Z, \) and \( U \) such that \( x \in X, z \in Z, \) and \( u \in U \). These constraints are either due to the equipment limitations, e.g., the allowed minimum/maximum mass flow rates, or the constraints to the operation decisions, e.g., the binary constraints for the chiller and TES operation decisions.

III. MPC PROBLEM

A. Problem Formulation

The objective of the optimization is to obtain a real-time optimal scheduling strategy for a chiller bank and TES to minimize the total electricity cost while satisfying the campus cooling demands and system constraints. We formulate an MPC problem to search for the optimal set points of the scheduling strategy.

The main idea of MPC is to use the system model of the central plant to predict its future evolution and obtain its optimal control on a moving horizon basis.

At each time step \( k \), we are interested in solving the following finite-horizon optimization problem regarding the state and control actions at time steps \( \{k, k+1, \ldots, k+N-1\} \), where \( N \) is the prediction horizon of the MPC problem:

\[
\begin{aligned}
\min \quad & \sum_{k=1}^{N-1} c(l|k) z_{1,l}(k \Delta t) \\
\text{s.t.} \quad & x(l+1|k) = f(x(l|k), z(l|k), u(l|k)) \\
& g_1(x(l|k), z(l|k), u(l|k), d(l|k)) = 0 \\
& g_2(z(l|k), u(l|k)) = 0, \quad l = k, \ldots, k+N-1 \\
& x(l|k) \in X, \quad z(l|k) \in Z, \quad u(l|k) \in U \\
& x(k|k) = x(k), \quad x(k+N|k) \in X_f
\end{aligned}
\]  

(14)

where \( x(l|k) \) denotes the state variable for the time step \( l \) predicted at time step \( k \) (similar notations also apply to other variables); \( X_k = \{x(l|k) | l = k, \ldots, k+N\}, Z_k = \{z(l|k) | l = k, \ldots, k+N-1\}, \) and \( U_k = \{u(l|k) | l = k, \ldots, k+N-1\} \) denote the predicted state, intermediate, and control sequences over the predicted horizon; electricity rate \( c(l|k) \) is assumed to be known for all time and \( z_1 = W_r \) denotes the electricity power.
consumed by the chiller bank; the cooling load $d(\|k\|$ is also the known input from the forecasting model described in the Appendix, i.e., $d(\|k\|) = Q(l\|k\|); x(\|k\|)$ denotes the actual state of the TES at time $k$; and $X_f$ denotes the terminal constraint set for the TES state, which is chosen to ensure stability and feasibility of the MPC problem [12], [18].

By assuming that the maximum campus cooling load never exceeds the maximum capacity of a chiller bank, there always exist feasible solutions for problem (14). Among all feasible solutions, let $(X_{k}^*, Z_{k}^*, U_{k}^*)$ denote the optimal one for problem (14). At time step $k$, only the first element of $U_{k}^*$ is implemented to the system, i.e., $u(k) = u^*(k\|k\|)$. At time step $k + 1$, the horizon of interests moves one-step forward to $(k + 1, \ldots, k + N)$, and the optimization problem (14) is solved again by changing index from $k$ to $k + 1$. The initial point of the predicted state sequence is taken as $x(k+1\|k\|+1) = x(k+1)$, where $x(k+1)$ denotes the newly measured state at time step $k+1$. This procedure is repeatedly implemented by moving forward the horizon step by step, which yields a moving horizon control strategy.

B. Conventional Approaches Based on MINLP

For the MPC optimization problem (14), at each time-step $k$, there are in total $N \ast (4n_c + 8)$ decision variables to be computed, of which there are $N \ast (n_c + 1)$ binary on, and $N \ast (3n_c + 7)$ real on. We also note that the constraints due to (13) have nonconvex terms due to the bilinear system dynamics. Thus, the optimization problem (14) is a nonconvex MINLP problem, which is extremely difficult to solve [22].

There are many software packages that solve nonconvex MINLP problems to proven optimality, such as BARON, α-BB, and Couenne [22]. But they require many computing resources and a large amount of computation time to achieve a certain global optimality. On the other hand, many convex-MINLP solvers can provide heuristic solutions to nonconvex MINLP in a more efficient way, e.g., BONMIN, DICOPT, and LaGO [24]. But they suffer from the high sensitivity to the initial starting point. All of these solvers can therefore be used off-line to compute the optimal set points but not suitable for real-time implementation of an MPC scheme.

C. Proposed Approach Based on DP and MILP

In this paper, we propose a heuristic algorithm to approach the nonconvex problem (14) based on DP and MILP. Our idea is to explore the special structure of the original problem and solve a set of subproblems with less complexity. The goal is to obtain suboptimal solutions to (14) in a time-efficient manner with certain performance guarantee. The proposed algorithm is described below.

- **Generating optimal TES operation profiles using DP**: We search the optimal TES and chiller operations in two steps. In the first step, we generate an optimal TES operation profile for a fixed future horizon by using DP. The system dynamics is approximated by a finite state machine (FSM) to reduce the computational complexity. Then, we fix the TES operation profile for the fixed future horizon, and solve for the optimal chiller operations by using a reformulated MPC problem.

- **Linearizing system dynamics**: The system dynamics is linearized along the predicted trajectory of state and control, where the trajectory is generated by simulating the nonlinear system model. The linearized dynamics can well approximate nonlinear one around the operating points.

- **Obtaining optimal chiller operations by solving an MILP problem**: In the second step, we reformulate the MPC problem with the linearized system dynamics and fixed TES operation profile, which leads to an MILP problem at each time step. It can be solved very efficiently with certain performance guarantee.

- **Generating nominal trajectories for linearization**: By simulating the nonlinear system model with the predicted control sequence obtained from MILP, we obtain the predicted trajectories of state and control for the next horizon. They are then used as nominal values in the next time step to linearize the system dynamics. The conceptual framework of the proposed approach is shown in Fig. 2 with its details to be provided next.

IV. DP-BASED ALGORITHM FOR OPTIMAL TES OPERATION PROFILE

Here, we describe a DP-based algorithm that generates an optimal TES operation profile $\sigma^{dp}$. In Section V, $\sigma^{dp}$ will be fixed in the MPC problem (14) for a fixed future horizon to solve for the optimal chiller operations. The basic idea is to quantize the state variable pair $(T_a, T_b)$ and then approximate the differential equation-based TES model with an FSM. Then we can obtain a sub-optimal TES operation profile with drastically lower computational complexity.

A. Approximating TES Dynamics by FSM

An FSM consists of a set of states including an initial state, an input alphabet, and a transition function that maps input symbols and current states to the next ones [25]. We use the following notation to define it: $FSM := (S, U, F, s_0)$, where $S$ denotes the set of states, $U$ denotes the input alphabet, also termed as “actions,” $F : S \times U \rightarrow S$ denotes the transition function of the state dynamics, and $s_0 \in S$ denotes the initial state. If $s(k) \in S$ is the current state and $u(k) \in U$ is the current action signal at the $k$th time step, then the next state $s(k+1) = F(s(k), u(k))$.

To develop a simple FSM model to approximate the chiller and TES system dynamics, we impose the following assumption only in this section.

**Assumption 1**: Both chilled water supply temperature $T_{es}$ and return one $T_{ev}$ are constants.

In reality, Assumption 1 is approximately satisfied with small fluctuations on $T_{es}$ and $T_{ev}$. This assumption helps develop a simple FSM model to approximate chiller and TES dynamics.

The set of states $S$ in FSM is defined as the collection of discretized states of TES: $s := \{T_{a}^d, T_{b}^d \} \in S$. Recall that a TES state is denoted by $(T_a, T_b)$, where $T_a \in [T_{es}, T_{ev}]$ is the top layer water temperature and $T_b \in [T_{es}, T_{ev}]$ is the bottom layer water temperature. By partitioning the interval $T_{es}, T_{ev}$ uniformly into $n_{ad}$ levels, we obtain the discretized version of $T_{a}^d$ and $T_{b}^d$, respectively. We add a further constraint that $T_{a}^d < T_{b}^d$ to be consistent with the reality. After the discretization, the cardinality of set $S$ is $n_{ad}^2 \times 1/2$.
The action of FSM at time step $k$ is taken as $u(k) := b(k)$, a $n_c$-by-1 vector containing the ON/OFF set-points for $n_c$ chillers. The action set $U$ is the $n_c$-dimensional binary vector space with the cardinality $2^{n_c}$.

The transition function $F$ is a map that takes each pair $(s, u)$ to some state $s' \in S$. This map is obtained by approximating the chiller and TES dynamics given by (10) – (11). The procedure to obtain $F$ is described as follows: At time step $k$, given the action $u(k)$, variables $\bar{m}, \dot{Q}_c, W_c(k)$ are completely obtained via (2) – (7) under Assumption 1. Note that the campus cooling load $\dot{Q}(k)$ is a known exogenous input predicted from a linear regression model (Appendix A). The TES operating mode $\sigma$ is then obtained as

$$\sigma(k) = \begin{cases} 1, & \text{if } \dot{Q}_c(k) \geq \dot{Q}(k) \\ 0, & \text{if } \dot{Q}_c(k) < \dot{Q}(k) \end{cases}$$

Given the current FSM state $(T_a^d(k), T_k^d(k))$, we then compute variables $T_a(k), T_k(k), \bar{m}, \dot{m}, \dot{Q}_c(k)$. If $\sigma(k) = 1$, i.e., TES is being charged, we compute these variables as follows:

$$T_a = T_{ca}$$
$$\bar{m} = \bar{m}_c + \bar{m}_t$$
$$T_k = \frac{\dot{m}_t T_{cr} + \bar{m} T_a^d}{\dot{m}_t}$$

If $\sigma(k) = 0$, i.e., TES is being discharged, we compute

$$T_a = T_{ce}$$
$$\bar{m} = \bar{m}_c + \bar{m}_t$$
$$T_k = \frac{\dot{m}_t T_{ce} + \bar{m} T_k^d}{\dot{m}_t}$$

The next FSM state $(T_a^d(k+1), T_k^d(k+1))$ is obtained by evolving discretized TES dynamics (10d)–(10e) or (11d)–(11e) one time-step forward with the above computed variables.

**Algorithm 1: DP for Generating Optimal State-Action Pairs**

**Input:** Input: $s^*, M_b, \dot{Q}(l)$, for $i = k, \ldots, k + N - 1$.

**Output:** $V_0(s, l), \forall s \in S$, $U^*(s, l)$, for $s \in S$ and $l = k, \ldots, k + N - 1$

1: Initialize cost function: $V(s, k + N) \leftarrow \Psi(s - s^*), \forall s \in S$.
2: Initialize feasible state set: $S_{k+N} \leftarrow s^*$.
3: Initialize index: $l \leftarrow k + N - 1$.
4: while $l \geq k$ do
5: $S_l \leftarrow \emptyset$.
6: for all $s \in S$ do
7: $V_{min} \leftarrow M_b$.
8: for all $u \in U$ do
9: $\hat{s} \leftarrow F(s, u)$.
10: if $\hat{s} \in S_{l+1}$ and $\hat{s}, s, u, \dot{Q}(l)$ satisfy constraints then
11: add $s$ to $S_l$.
12: $V \leftarrow c([l]c(u)) + V(\hat{s}, i + 1)$.
13: if $V < V_{min}$ then
14: $V(s, l) \leftarrow V, U^*(s, l) \leftarrow u, N S(s, l) \leftarrow \hat{s}, V_{min} \leftarrow V$.
15: end if
16: end if
B. Optimal TES Operation Profile Generated Through DP

At time step $k$, the electricity cost is given by
\[ c(k)e(u(k)) \]
where $c(k)$ denotes the electricity rate and $e(u(k))$ denotes the total electricity consumed by the chiller bank. A terminal cost $\Psi(s - s^*)$ is introduced to penalize the deviation from desired terminal state $s^*$, where function $\Psi(.)$ is quadratic.

We consider a finite-horizon optimization problem:

1) Problem 1: At time step $k$, given current state $s(k)$, optimization horizon $N$, and desired terminal state $s^*(k+N)$, compute the optimal sequence of $N$ actions $u(k), \ldots, u(k+N-1)$ such that, along the resulting state trajectory $s(k+1), \ldots, s(k+N)$, the campus cooling demands $\dot{Q}(k), \ldots, \dot{Q}(k+N-1)$ are met, and the finite-horizon cost $V^{opt} := \Psi(s(k+N) - s^*) + \sum_{k=0}^{N-1} c(k)e(u(k))$ is minimized.

Note that the campus cooling demands should be strictly satisfied throughout optimization horizon $N$. This constraint implies that the set of acceptable control actions is time-varying. We denote this set of acceptable control actions at time step $k$ as $U_k$. In fact, $U_k$ depends on both $Q(k)$ and $s(k)$. Similarly, not all states $s(k+1)$ can be reached from arbitrary given previous states $s(k)$ with action $u(k) \in U_k$. Problem 1 is solved via DP as presented in Algorithm 1. At time step $k$, Algorithm 1 generates the following outputs: 1) the optimal action $U^*(s,l)$; 2) the next state $N[ss]$, obtained through FSM transition with action $U^*(s,l)$; and 3) the minimal cost $V(s,l)$ taking the system from state $s$ at time $l$ to final state $s^*$ at time $k+N$, where $s \in S$ and time $l = k, \ldots, k+N-1$.

Algorithm 1 starts from a terminal state, and searches the optimal state trajectory and corresponding actions backwards in time. We use $S_l$ to denote the “feasible state set” at time step $l$, which is the set of states that can be brought to the desired terminal state $s^*$ at time $k+N$ by a sequence of actions. The algorithm begins with setting $S_{k+N}$ to $s^*$. The cost $V(s,l)$ at time $k+N$ is set to be the terminal cost $\Psi(s - s^*)$. At each time step $l$, all possible states $s \in S$ and all possible actions $u \in U$ are evaluated through transition function $F$. Those states that can be brought to $S_{l+1}$ while satisfying all constraints are marked as “feasible” ones. The most cost-effective action at state $s$ is found and recorded in $U^*(s,l)$. The corresponding cost value is computed and stored in $V(s,l)$. Based on the outputs generated from Algorithm 1, the optimal TES operation profile $\{s^d(l)\} - k, \ldots, k+N-1$ can be easily derived. This is achieved by traversing the optimal state-action pairs $\{N[ss], U^*(s,l), V(s,l) s \in S, l = k, \ldots, k+N-1\}$ forward in time, starting from initial state $s(k)$. This procedure is summarized in Algorithm 2.

Algorithm 2: Generate TES Operation Profile from DP result

One major advantage of a DP-based approach to choose a TES operation profile is that, unlike in [18], [19] where the TES operation profiles are fixed as a priori, they are decided by an optimization algorithm by taking the predicted campus cooling load into consideration. The complexity of the DP-based algorithm depends on the number of FSM quantization levels $n_d$ and the number of chillers $n_c$, i.e., the algorithmic complexity is of the order $O(N \star 2^{n_c} \star (n_d(2n_d + 1)/2))$. Note that the number of chillers $n_c$ and the length of horizon $N$ are fixed here. With more quantization levels, we can have a better approximation of TES dynamics. By varying the number of quantization levels $n_d$, we can make a tradeoff between its complexity and accuracy to approximate the TES dynamics.

V. DECIDING OPTIMAL CHILLER OPERATIONS

Here, we present an MILP-based method to optimally operate the chiller bank in real time. Its main idea is to reformulate an MPC problem with linearized system dynamics and a fixed TES operation profile. The latter is generated via the DP algorithm. This formulation leads to an MILP problem, which can be solved very efficiently at every time step.

A. Linearizing System Dynamics

It appears that the major difficulties of solving a nonconvex MINLP problem are due to their non-convexity, which prevents the implementation of efficient algorithms [22]. For problem (14), nonconvexity mainly comes from the bilinear terms in the dynamics of the chiller bank and TES.
Let \( \tilde{x}(k) \), \( \tilde{z}(k) \), and \( \overline{u}(k) \) denote the nominal values of the state and control at the \( k \)th time step. We approximate the bilinear dynamic equation (13a) by a set of linear dynamic equations

\[
x[k+1] = A_1(k)x(k) + A_2(k)z(k) + B(k)u(k) + E_0(k)
\]

where matrices \( A_1(k) \), \( A_2(k) \), \( B(k) \), and \( E_0(k) \) are obtained through linearization of (13a) around nominal values \( \tilde{x}(k) \), \( \tilde{z}(k) \), and \( \overline{u}(k) \). Similarly, we approximate bilinear balance equation (13b) by a set of linear balance equations

\[
C_{11}(k)x(k) + C_{12}(k)z(k) + D_{11}(k)u(k) + D_{12}(k)d(k) = P_1(k).
\]

To remove the bilinear terms introduced by the product of a binary valued variable and a real valued function, we employ an equivalence modeling trick from [21]:

1) Property 1: Let \( \xi(\omega) \) be a real valued scalar function defined on a bounded set \( \Omega \) such that \( m \leq \xi(\omega) \leq M \) for any \( \omega \in \Omega \), \( \kappa \in [0, 1] \) be a binary valued scalar variable, and \( \mu \in \mathbb{R} \) be a real valued scalar variable. For any \( \omega \in \Omega \), we have the following equivalence:

\[
\mu - \kappa \xi(\omega) \iff \begin{cases} 
\mu \leq M \kappa \\
\mu \geq \mu \kappa \\
\mu \leq \xi(\omega) - m(1 - \kappa) \\
\mu \geq \xi(\omega) - M(1 - \kappa).
\end{cases}
\]

Note that all bilinear terms appearing in (13c) are in the form of a binary valued variable and a real valued function. We apply Property 1 to convert the bilinear equation (13c) equivalently to a set of linear inequalities

\[
C_{21}z(k) + D_{21}u(k) \leq P_2(k).
\]

### B. Formulating an MILP Problem

Next, we define a simplified MPC optimization problem with a linearized model. The TES operation profile is fixed as a known input \( \sigma^{\text{fix}} \) to reduce the computational complexity. We fix the TES operation profile as \( \sigma^{\text{fix}} = \sigma^{\text{op}} \), where the TES operation profile \( \sigma^{\text{op}} \) is generated every \( N \) time steps using DP as described in Section IV. At step time \( k \), if \( \text{mod}(k, N) = 0 \), we set \( k_N = k \) and generate \( \{\sigma^{\text{op}}(k), \ldots, \sigma^{\text{op}}(k + N - 1)\} \) via Algorithm 1; otherwise, if \( \text{mod}(k, N) \neq 0 \), we use \( \sigma^{\text{op}} \) generated before. In the simplified MPC formulation, \( \sigma^{\text{fix}} \) is then obtained at each step time \( k \) as

\[
\sigma^{\text{fix}}(l,k) = \sigma^{\text{op}}(k + \text{mod}(l, N)), \quad l = k, \ldots, k + N - 1.
\]

At each step time \( k \), we choose a nominal trajectory of state and control \( \{\tilde{x}(l,k), \tilde{z}(l,k), \overline{u}(l,k)\} | l = k, \ldots, k + N - 1 \). The linearized system dynamics for the predicted horizon \( k, \ldots, k + N - 1 \) is obtained along the nominal trajectory of the state and control. Then, we consider the following optimization problem as an approximation to problem (14):

\[
\min_{x_k, z_k, u_k} \sum_{i=k}^{k+N-1} c(l[k])z_{1 i k} \Delta t
\]

s.t. \( x[l + 1 k] = A_1(l[k])x[l k] + A_2(l[k])z[l k] + B(l[k])u[l k] + E_0(l[k]) \)
\( C_{11}(l[k])x[l k] + C_{12}(l[k])z[l k] + D_{11}(l[k])u[l k] + D_{12}(l[k])d[l k] = E_2(l[k]) \)
\( x[l k] \in \mathcal{X}, \quad z[l k] \in \mathcal{Z}, \quad u[l k] \in \mathcal{U} \)
\( u_1(l[k]) = \sigma^{\text{fix}}(l[k]), \quad l = k, \ldots, k + N - 1 \)
\( x[k k] = x[k], \quad x[k + N k] \in \mathcal{X}_f \).

Problem (15) is an MILP problem, with \( N \times n_u \) binary variables and \( N \times (3n_T + 7) \) real on. Due to its convexity, it is much easier to solve than the nonconvex MINLP problem in general. Many software packages can efficiently solve an MILP problem with certain performance guarantee, e.g., SCIP, GLPK, IpXsolve, and CBC [24, 26].

### C. Generating Nominal Trajectories for Linearization

Recall that, to solve problem (15) at each time step \( k \), we require a nominal trajectory of the state and control for the predicted horizon \( \{k, k + 1, \ldots, k + N - 1\} \). In this paper, we generate it through the simulation of the nonlinear system model (13) with optimal predicted control computed from problem (15). Let \( \{X^*_k, Z^*_k, U^*_k\} \) denote the optimal solution to (15) at time step \( k \). For the next predicted horizon \( \{k + 1, \ldots, k + N\} \), we heuristically choose the following control sequence as the nominal trajectory of control:

\[
\overline{u}[l[k + 1]] = \begin{cases} 
\overline{u}^*(l[k]), & \text{if } l = k + 1, \ldots, k + N - 1 \\
\overline{u}^*(k[k]), & \text{if } l = k + N.
\end{cases}
\]

Then, we apply the control sequence \( \{\overline{u}[l[k + 1]] | l = k + 1, \ldots, k + N\} \) to simulate the nonlinear system model (13) with the predicted load input \( \{d[l[k + 1]] | l = k + 1, \ldots, k + N\} \) and initial state \( \overline{x}[k + 1] = \overline{x}[k + 1|k + 1] \). The simulated state sequences \( \{\overline{x}[l[k + 1]] | l = k + 1, \ldots, k + N\} \) and \( \{\overline{z}[l[k + 1]] | l = k + 1, \ldots, k + N\} \) are taken as the nominal trajectory of states for the predicted horizon \( \{k + 1, \ldots, k + N\} \). Note that the initial trajectory of the state and control at time step 0 can be generated with any control sequence based on the experience or other heuristics.

The MILP-based algorithm is described in Algorithm 3 to compute suboptimal system control set-points of the chiller bank. Note that the TES operation profile is determined via Algorithm 1 for every horizon.

**Algorithm 3.** MILP-based Algorithm for Generating Optimal System Set-points of the Chiller Bank

**Input:** System parameters, optimization horizon \( N \), sampling time \( \Delta t \), total time steps \( K \), and performance gap \( \varepsilon \).

**Output:** Optimal system control set points \( \{u(k)\}_{k=0}^{K} \).
1: Initialize nominal values \( \{x(l), \tilde{x}(l), \tilde{u}(l)\}_{l=0}^{N-1} \) and set time step \( k = 0 \).

2: repeat

3: Measure the current state variable \( x(k) = [T_a(k), T_b(k)] \).

4: Predict campus cooling loads \( \{d(l \mid k)\}_{l=k}^{N-1} \) by (17).

5: if \( m \mod (k, N) = 0 \) then

6: Set the current horizon starting index \( k_N = k \).

7: Generate the TES operation profile \( \{\sigma(l)\}_{l=k}^{N-1} \) via Algorithms 2 and 1.

8: end if

9: Fix TES operation profile as known control input \( \{\sigma(l)\}_{l=k}^{N-1} = \sigma^* \).

10: Linearize system model (13) from time step \( k \) to \( k + N - 1 \) around nominal values \( \{x(l \mid k), \tilde{x}(l \mid k), \tilde{u}(l \mid k)\}_{l=k}^{N-1} \).

11: Compute \( \{x^*(l \mid k), z^*(l \mid k), u^*(l \mid k)\}_{l=k}^{N-1} \) by solving MILP problem (15) with performance gap \( \epsilon \) and fixed TES operation profile \( \sigma^* \).

12: Determine the next horizon nominal control sequence \( \{\tilde{u}(l + 1)\}_{l=k+1}^{K} \) by (16).

13: Obtain \( \{x(l + 1), \tilde{x}(l + 1)\}_{l=k+1}^{K} \) by simulating the system model (13) with \( \{\tilde{u}(l + 1)\}_{l=k+1}^{K} \).

14: Set time step \( k + 1 \rightarrow k \).

15: until The final time step \( k = K \).

16: return System control set points \( \{u(k) = u^*(k \mid k)\}_{k=0}^{K} \).

VI. CASE STUDY

Here, we present a case study to demonstrate MPC of a central chiller plant in summer time at the University of California, Irvine (UCI). We first use the measured data from the UCI campus to calibrate the parameters of the system model described in Section II. Then, the system model is built up in the MATLAB environment. An MPC strategy is compared with a baseline strategy through MATLAB simulations.

A. System Setup

The UCI central plant is depicted in Fig. 1. A bank of \( n_c = 7 \) chillers supplies chilled water to satisfy the campus cooling demand. The supply water temperature of all chillers is the same and maintained at a constant value \( T_{cw} = 40 \degree F \). The return water temperature \( T_{cr} \) is the same for all chillers, and varies according to the system dynamics. For the MPC strategy considered below, we constrain the return water temperature \( T_{cr} \) in the range of \([53, 63]\)\degree F to ensure all chillers to work with high efficiency. They can only work in an ON/OFF mode, either with the maximum or zero mass flow rate. The maximum ones are fixed constant values and summarized in Table I. The parameters of the chiller electricity model (3) are calibrated through the regression analysis of the available measured data, which are also summarized in Table I. We calculate COP of each chiller using model (4) with parameters given in Table I. The calculated COP is depicted in Fig. 3 versus return water temperature in the range of interests. Intuitively, we observe that the seventh chiller has the highest efficiency and the fifth chiller has the lowest one in most range of return water temperature.

Chilled water supplied by the chiller bank can also be stored in a water tank TES for shifting the campus cooling load. The cross-layer area of TES is \( A = 564.95 \text{ m}^2 \) with a height of \( h = 30.48 \text{ m} \). At the initial time, TES is assumed to be fully discharged with the uniform water temperature of \( 60 \degree F \). The parameters of the TES model in (10) and (11) are calibrated through the regression analysis of the available measured TES data. The calibrated time-constant multipliers are \( f_{a,c} = 1.740, f_{b,c} = 2.222, f_{a,d} = 4.138, \) and \( f_{b,d} = 1.599 \). The calibrated inter-layer overall heat transfer coefficients are \( U_c = 0.0283 \text{ kW/(m}^2 \cdot \text{K}) \) and \( U_d = 0.0197 \text{ kW/(m}^2 \cdot \text{K}) \).

![Fig. 3. Plot of chiller COP for different return water temperature.](image)

| Table I | CALIBRATED PARAMETERS FOR CHILLER MODEL |
|---------|----------------------------------------|
| Chillers | \( m_{ch,i} (\text{kg/s}) \) | \( a_i (\text{kW/K}) \) | \( b_i (\text{kW}) \) |
| 1       | 119.87                               | 54.67                     | 318.3                     |
| 2       | 119.87                               | 54.67                     | 318.3                     |
| 3       | 138.80                               | 80.17                     | 225.0                     |
| 4       | 283.91                               | 143.6                     | 666.1                     |
| 5       | 226.59                               | 165.7                     | 592.8                     |
| 6       | 236.59                               | 109.2                     | 786.7                     |
| 7       | 283.91                               | 168.8                     | 332.1                     |

\(^1\text{Temperature } T \text{ can be converted as } T(K) = 273.15 + (T(\degree F) - 32)/9.\)
Fig. 4. Parameter training results of campus cooling load prediction model for the month of June 2009. The top plot compares predicted campus cooling load (blue solid line) and actual campus cooling load (red dotted line). The bottom plot depicts the corresponding ambient temperature (blue solid line) and ambient humidity ratio (green dashed line).

The campus cooling load is modeled and predicted using an ARX (autoregressive exogenous) model, where the exogenous inputs are taken as ambient temperature and humidity ratio (see the Appendix for more information). The parameters used in cooling load model (17) are trained using the actually measured campus load and weather data in the month of June 2009. The values of trained parameters are omitted here due to the space limitation, while the training results are depicted in Fig. 4. We observe that the actual cooling loads are well approximated by the cooling load prediction model (17). In the MPC problem, we employ the trained cooling load model (17) to predict the future cooling load based on historical campus cooling load data and weather forecast data. The electricity rate schedule for summer hours are obtained from Pacific Gas and Electric Company and summarized in Table II, for simplicity only the Time-Of-Use (TOU) energy charge is used in the cost function and the demand charge part is not included.

| TABLE II | ELECTRICITY RATES FOR SUMMER HOURS |
|----------|-----------------------------------|
| Summer   | Period   | Price ($/kWh) |
| On-Peak  | 12:00-18:00 | 0.1397        |
| Part-Peak| 08:30-12:00, 18:00-21:30 | 0.0998        |
| Off-Peak | 21:30-08:30 | 0.0835        |

This suggests that TES should be charged for the first 9 hours and switched to a discharge mode thereafter.

1) **Baseline:** This strategy generates suboptimal chiller and TES operations based on a greedy search algorithm. All possible 27 chiller ON/OFF modes are first ordered according to their efficiency. Then, under a given TES charging/discharging profile, an initial feasible solution is obtained which requires more than necessary chillers staged ON. Starting from this initial solution, a greedy algorithm is applied to search for better chiller operation sequences that most drastically reduce the TOU-based cost function, until some given constraints are violated. This strategy has been tested with historical operation data and is shown to achieve 30% saving over the plant operator’s heuristic operation in [19].

2) **MPC:** For the MPC strategy, we first consider the MINLP approach to solve the MPC problem (14) with 192 binary variables and 692 real variables at every time step. Due to the large

B. Two Scheduling Strategies

Here, we consider two strategies to schedule the operation of TES and chiller bank. One is a greedy-search-based heuristic approach which serves as a baseline strategy, and the other is the proposed DP and MILP-based MPC strategy. To simplify the comparison study, in both strategies, the TES charging/discharging operation profile is generated by using Algorithms 1 and 2 with horizon \( N = 24 \) and desired terminal state \( s^* = (58 \, ^\circ F, 58 \, ^\circ F) \). The system dynamics are approximated by an FSM quantized with \( \pi_d = 20 \) levels. With the campus cooling load shown in Fig. 5(a), the resulting TES operation profile \( \delta^{\text{dy}}(k) \) is given as

\[
\delta^{\text{dy}}(k) = \begin{cases} 
1, & \text{if } 1 \leq k \leq 9 \\
0, & \text{if } 10 \leq k \leq 24.
\end{cases}
\]

This suggests that TES should be charged for the first 9 hours and switched to a discharge mode thereafter.

1) **Baseline:** This strategy generates suboptimal chiller and TES operations based on a greedy search algorithm. All possible 27 chiller ON/OFF modes are first ordered according to their efficiency. Then, under a given TES charging/discharging profile, an initial feasible solution is obtained which requires more than necessary chillers staged ON. Starting from this initial solution, a greedy algorithm is applied to search for better chiller operation sequences that most drastically reduce the TOU-based cost function, until some given constraints are violated. This strategy has been tested with historical operation data and is shown to achieve 30% saving over the plant operator’s heuristic operation in [19].

2) **MPC:** For the MPC strategy, we first consider the MINLP approach to solve the MPC problem (14) with 192 binary variables and 692 real variables at every time step. Due to the large
number of variables and many nonlinear equations, the conventional nonconvex MINLP solvers (e.g., Couenne) cannot provide any feasible solution or does not converge in a reasonable time frame (up to 5 h). The convex-MINLP solvers (e.g., Bonmin) can always heuristically generate some feasible solutions in around 20 min but always keep running without stopping up to 5 h (this means it is still searching for better solutions). For a comparison purpose, we record those solutions obtained from software Bonmin by stopping it when the computation time reaches one hour.

Then, we consider the proposed approach for solving the MPC problem based on DP and MILP. As discussed above, the optimal TES operation profile $\delta_0^p$ is generated using DP. By fixing $\delta_0^p$ as a known input, we formulate an MILP-based MPC problem (15) to solve for optimal chiller operations. The MILP-based heuristical Algorithm 3 is employed to search for the optimal set points of the system. The sampling time $\Delta t = 1$ h is used to discretize the system model. The prediction horizon of the MPC scheme is taken as $N = 24$ h. At each hour, the software package SCIP is employed to solve the MILP problem (15) with performance gap $\epsilon = 5\%$. For the MILP problem considered in this paper, it usually takes less than 10 minutes only to achieve the performance requirement. The initial nominal trajectory used by Algorithm 3 is obtained by simulating the system with the baseline strategy. By comparing the solutions obtained using two approaches, it is found that the solutions obtained through DP and MILP are always better than the solutions obtained by convex-MINLP approach. In Section VI-C, we will compare the baseline and MPC strategies with the solution computed via DP and MILP approaches.

C. Comparison of Two Strategies

The performance of both strategies is compared through MATLAB simulations with historical weather data for a specific day in June 2010. The total simulation time is $K = 24$ h. The predicted campus cooling load profile is depicted in Fig. 5(a) overlaid with the electricity rate for summer hours. The predicted campus cooling load profile is depicted in Fig. 5(a) overlaid with the electricity rate for summer hours. We duplicate the campus cooling load profile for the next day to simulate the MPC scheme.

The controls computed from two strategies are applied to the simulation model of the UCI central plant. The chilled water temperature through the campus and the chiller bank are depicted in Fig. 6(a) and (b), respectively. The top and bottom temperatures of TES are depicted in Fig. 6(c). We observe that chilled water temperatures in the MPC strategy have larger variations than those in the baseline strategy. We also observe that for the MPC strategy the top and bottom layers of TES reach the almost same temperature at the end of day. But for the baseline strategy the bottom layer has much lower temperature than...
TABLE III

| Strategy | Electricity consumption (kWh) | Electricity cost ($) |
|----------|-------------------------------|----------------------|
| Baseline | 145,746                        | 15,040               |
| MPC      | 131,616                        | 13,411               |

the top layer of TES, indicating that there is additional unused cooling energy stored in TES.

For both strategies, the hourly cooling amount provided by individual chillers and TES are illustrated in Fig. 7, where the negative cooling amount means that TES is in charging mode. We observe that in both strategies chiller 5 is never operated, and chillers 1, 2, and 7 are operated in most of time. This makes sense since chiller 5 has the lowest COP while chillers 1, 2, and 7 have almost highest COP in the range of interests for the chilled water return temperature as shown in Fig. 3. However, two strategies are different in the way to charge/discharge TES. We observe that TES charges more cooling amount in the baseline strategy than in MPC during the off-peak hours. However, TES discharges less cooling in the baseline strategy than in MPC during the on-peak hours. The overcharge in off-peak hours and insufficiently discharge in on-peak hours are the main causes of energy inefficiency under the baseline strategy. The energy and economical inefficiency can also be observed from electricity consumption and cost plot in Fig. 5(b) and (c). We summarize the total electricity consumption and cost in Table III. From it, the MPC-based strategy performs better than the baseline by around 9.70%, in terms of total electricity consumption, and by around 10.84%, in terms of total electricity cost.

D. Discussion

1) Horizon Length: The horizon length has an important impact on the performance of solutions and the computation time of algorithms. For the case study presented in this paper, we choose the horizon length as $N = 24$ h for the illustration of simulation results. We have also tested different horizon values for a comparison purpose. Generally speaking, by choosing a longer horizon length (e.g., multiples of 24 h), we obtain optimal solutions that take into consideration of the forecast information and system dynamics in a long time range. But the computation time for obtaining the solution increases a great deal as the length of horizon increases. On the other hand, by choosing a shorter horizon length, we obtain the solutions with much less computation time. But the solutions may be only optimal for a short time range and the performance of the MPC strategy can be largely degraded. In practice, one can try a couple of horizon lengths and choose the one that satisfies the performance and computation time requirements.

2) Uncertainty Analysis: The simulation results presented here are obtained under the assumption that the parameters are accurate and chiller system models can perfectly represent the real system. However, in reality, there usually exist uncertainties in parameters and models. In this section, a simple Monte Carlo uncertainty analysis is carried out to demonstrate how the modeling uncertainty would affect the potential energy saving of proposed MPC strategies [27]. Here, we consider the uncertainties for some major parameters of the chiller system, namely, $T_{ca}$ (chilled water supply temperature), and $f_{ac}$, $f_{bc}$, $f_{ad}$, and $f_{bd}$, $U_c$ and $U_d$ (fitted parameters of the TES model). We assume the considered parameters all follow the normal distribu-
tions with means equal to their nominal values and standard deviations equal to 5% of their nominal values. A sample matrix $M_{k}$ is generated based on random sampling of the considered parameters, where $n$ denotes the number of parameters and $k$ denotes the sample size. Here we have $n = 7$ parameters and we take $k = 5060$ as the sample size. Then we repeatedly apply baseline and MPC strategies using the randomly generated parameter samples. The probabilistic distributions of energy savings (electricity consumption or cost) can be obtained via the numerical approximations. For the case study considered here, the standard deviations of energy savings are less than 2.3% compared with around 10% energy savings for nominal values. The results show that the fluctuation of energy savings caused by parameter uncertainties is relatively small. Thus, the proposed MPC strategy is not very sensitive to the modeling parameters specified here.

VII. CONCLUSION

This work formulates an MPC problem to obtain optimal set-points for supervisory control of a central chiller plant with thermal energy storage (TES). It proposes an efficient heuristic algorithm to solve the MPC problem based on dynamic programming (DP) and mixed integer linear programming (MILP). The optimal TES operation profile is obtained for a fixed horizon by solving a DP problem, and then optimal chiller operations are obtained at each time step by solving an MILP problem with a fixed TES operation profile. Simulation results show that both electricity consumption and cost can be significantly reduced for the campus central plant at the University of California, Irvine. Future research includes more benchmark studies with other control methods [28], [29] and using the optimization methods successfully used in other fields to solve this problem [30]–[49].

APPENDIX

CAMPUS COOLING LOAD PREDICTION

For the MPC problem, we need to know the predicted campus cooling load for the next $N$ hours, where $N$ is the prediction horizon. In this paper, we employ a linear ARX model to predict the campus cooling loads [50]. Mathematically, at each time step $k$, we predict the campus cooling loads for the next $N$ time steps. For $l = k, \ldots, k + N - 1$, we have

$$Q(l | k) = \sum_{i=1}^{N} \alpha_i Q(l-i | k) + \sum_{i=0}^{N} \beta_i T_{am}(l-i | k) + \gamma_i h_{am}(l-i | k)$$

where $Q(l | k)$, $T_{am}(l | k)$, and $h_{am}(l | k)$ denote the predicted cooling loads, ambient temperature, and ambient humidity ratio for time step $l$ at the time step $k$, respectively. The parameters $\alpha_i$, $\beta_i$, and $\gamma_i$ can be obtained using historically measured campus cooling load and weather data. Note that here we only consider the ambient temperature and humidity ratio as the inputs to the ARX model, which have most significant correlation to the campus cooling loads. To further improve the prediction accuracy, one could take more variables as the inputs to the ARX model, e.g., ambient solar radiation and campus occupancy information.

The past values of campus cooling loads are available on an hourly basis from the campus plant management system. At time step $k$, we assume the knowledge of the actual values of $Q$ up to time step $k$: $\hat{Q}(l | k) = Q(l)$, for all $l \leq k$. The past information of ambient temperature $T_{am}$ and humidity ratio $h_{am}$ are available from on-site measurement. The predicted weather information (ambient temperature $T_{am}$ and ambient humidity ratio $h_{am}$ for next few days) can be obtained from National Digital Database via NOAA’s National Weather Service. At time step $k$, we assume the knowledge of the actual values $T_{am}$ and $h_{am}$ up to time step $k$, and the predicted values of $T_{am}$ and $h_{am}$ for future time steps:

$$T_{am}(l | k) = \begin{cases} T_{am}(l), & \text{if } l \leq k \\ \hat{T}_{am}(l), & \text{if } l > k \end{cases}$$

$$h_{am}(l | k) = \begin{cases} h_{am}(l), & \text{if } l \leq k \\ \hat{h}_{am}(l), & \text{if } l > k \end{cases}$$

REFERENCES

[1] U.S. Dept. Energy, “2011 Buildings Energy Data Book,” Mar. 2012 [Online]. Available: http://buildingsdatabook.eren.doe.gov
[2] “2011 ASHRAE Handbook–HVAC Applications,” ASHRAE Inc., Atlanta, GA, USA, 2011.
[3] L. B. Hyman and F. Bockmiller, “Primary chilled water loop retrofit,” ASHRAE J., vol. 42, no. 12, pp. 60, 62–64, 2000.
[4] P. Mancarella and G. Chicco, “Real-time demand response from energy shifting in distributed multi-generation,” IEEE Trans. Smart Grid, vol. 4, no. 4, pp. 1928–1938, Nov. 2013.
[5] K. Kapoor, “Modeling and optimization for energy efficient large scale cooling operation,” Ph.D. dissertation, Dept. Chem. Eng., Univ. of Texas at Austin, Austin, TX, USA, 2013.
[6] R. T. Olson and S. Lieberman, “Optimization of a chilled water plant using sequential quadratic programming,” Eng. Optimiz., vol. 15, pp. 171–191, 1990.
[7] E. A. Koeppe1, S. A. Klein, J. W. Mitchell, and B. A. Flake, “Optimal supervisory control of an absorption chiller system,” HVAC&R Res., vol. 1, pp. 325–342, 1995.
[8] S. Wang and Z. Ma, “Supervisory and optimal control of building HVAC systems: A review,” HVAC&R Res., vol. 14, pp. 3–32, 2008.
[9] A. Rakkes and M. S. Waring, “Using multiobjective optimizations to discover dynamic building ventilation strategies that can improve indoor air quality and reduce energy use,” Energy and Buildings, vol. 75, pp. 272–280, 2014.
[10] W. J. Cole, D. P. Morton, and T. F. Edgar, “Optimal electricity rate structures for peak demand reduction using economic model predictive control,” J. Process Control, vol. 24, no. 8, pp. 1311–1317, 2014.
[11] B. Sun, P. B. Luh, Q. Jia, Z. Jiang, F. Wang, and C. Song, “Building energy management: Integrated control of active and passive heating, cooling, lighting, shading, and ventilation systems,” IEEE Trans. Autom. Sci. Eng., vol. 10, no. 3, pp. 588–602, Jul. 2013.
[12] B. J. Rawlings and D. Q. Mayne, Model Predictive Control: Theory and Design, Madison, WI, USA: Nob Hill, 2009.
[13] J. E. Braun, S. A. Klein, and J. M. W. Beckman, “Methodologies for optimal control of chilled water systems without storage,” ASHRAE Trans., vol. 95, no. 1, pp. 652–662, 1989.
[14] J. E. Braun, “Reducing energy costs and peak electrical demand through optimal control of building thermal storage,” ASHRAE Trans., vol. 96, no. 2, pp. 876–880, 1990.
[15] N. N. Kota, J. M. House, J. S. Arora, and T. S. Smith, “Optimal control of HVAC systems using DDP and NLP techniques,” Optimal Control Applications & Methods, vol. 17, pp. 71–78, 1996.
[16] G. P. Henze, R. H. Dodier, and M. Krarti, “Development of a predictive optimal controller for thermal energy storage systems,” HVAC&R Res., vol. 3, no. 3, pp. 233–264, 1997.
DENG et al.: MPC OF CENTRAL CHILLER PLANT WITH THERMAL ENERGY STORAGE VIA DYNAMIC PROGRAMMING AND MILP

Yan Lu (M’04) received the B.S. and M.S degrees from Tsinghua University, Beijing, China, and the Ph.D. degree from Carnegie Mellon University, Pittsburgh, PA, USA, all in electrical engineering.

She is currently a Senior Research Scientist with the National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA. Before joining NIST, she is the head of Production Operation and Optimization Research Group, Siemens Corporation, Corporate Technology. Her research interests cover distributed agent-based control, energy optimization, with Siemens, she has led and successfully delivered more than 10 million dollars of government funded research projects in the areas of survivable control systems, energy management and smart grid automation. From 2002 to 2004, she was with Seagate Research Center and responsible for developing hard disk drive servo controllers.

Jack Brouwer received the B.S. and M.S. degrees from the University of California, Irvine, CA, USA, in 1989 and 1987, respectively, and the Ph.D. degree from Massachusetts Institute of Technology, Cambridge, MA, USA, in 1993, all in mechanical engineering.

He is currently an Associate Professor of mechanical and aerospace engineering and an Associate Director of the National Fuel Cell Research Center and Advanced Power and Energy Program, University of California, Irvine, CA, USA. His research interests include advanced and alternative energy systems development, electrochemical conversion devices and systems, dynamic simulation and control, turbulent reacting flows, chemical kinetics, and electrochemical reactions with concurrent heat, mass and momentum transfer.

Prashant G. Mehta (M’06) received the Ph.D. degree in applied mathematics from Cornell University, Ithaca, NY, USA, in 2004.

He is currently an Associate Professor with the Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign (UIUC), Urbana, IL, USA. Prior to joining UIUC, he was a Research Engineer with the United Technologies Research Center (UTRC). His research interests are at the intersection of dynamical systems and control theory, including mean-field games, nonlinear filtering and model reduction.

Prof. Mehta was the recipient of several awards including an Outstanding Achievement Award for his research contributions at UTRC, Best Paper awards with his students at UIUC, and teaching and advising honors at UIUC. He currently serves on the editorial boards of the Systems and Control Letters and the ASME Journal of Dynamic Systems, Measurement, and Control.

Meng Chu Zhou (S’88–M’90–SM’93–F’03) received the B.S. degree in control engineering from Nanjing University of Science and Technology, Nanjing, China, in 1983, the M.S. degree in automatic control from Beijing Institute of Technology, Beijing, China, in 1986, and the Ph.D. degree in computer and systems engineering from Rensselaer Polytechnic Institute, Troy, NY, USA, in 1990.

He joined the New Jersey Institute of Technology (NJIT), Newark, NJ, USA, in 1990, where he is now a Distinguished Professor of electrical and computer engineering. He has authored and coauthored over 550 publications including 11 books, over 260 journal papers, and 22 book chapters. His research interests are in Petri nets, sensor networks, web services, big data, semiconductor manufacturing, transportation, and energy systems.

Prof. Zhou is a Life Member of the Chinese Association for Science and Technology-USA and served as its President in 1999. He is a Fellow of IFAC and AAAS. He is the founding Editor of the IEEE Press Book Series on Systems Science and Engineering and an associate editor of the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, and the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS.

Amit Chakraborty (M’96) received the B.Tech. and M.Tech. degrees in electronics and electrical communication engineering from the Indian Institute of Technology, Kharagpur, India, in 1988 and 1990, respectively, and the M.S., M.Phil., and Ph.D. degrees in electrical engineering from Yale University, New Haven, CT, USA, in 1991, 1993, and 1996, respectively.

He joined Siemens Corporate Research right after completing his doctoral studies. Currently, he is the head of Predictive Analytics research group within the Business Analytics & Monitoring Technology Field at Siemens Corporation, Corporate Technology, Princeton, NJ, USA. He leads the predictive analytics team that focuses primarily on Data Mining and Optimization applications across all Siemens business units. He has successfully led several research and development projects in energy management such as total plant optimization for high performance buildings, parametric fault diagnostics for gas turbines, power flow optimization for electric grids and semantic data modeling for power systems, etc. He holds 20 patents with many more pending and has authored and coauthored over 50 publications in respected journals and conferences.