

# An approximate dynamic programming approach to the energy management of a building cooling system

Nicola Ceriani, Riccardo Vignali, Luigi Piroddi, Maria Prandini

**Abstract**—This paper is concerned with optimal energy management of micro-grids. The goal is to show that the problem of minimizing the operating costs of a micro-grid by coordinating and scheduling its components can be formulated as a constrained optimal control problem for a stochastic hybrid system. This, in turn can be addressed through the Dynamic Programming (DP) approach, and the resulting DP equations solved through approximate DP techniques. A simple case study of a building cooling system with two chillers serving a cooling load is presented to this purpose.

## I. INTRODUCTION

Optimal energy management of micro-grids has gained considerable attention in recent years (see, *e.g.*, [1] [2], [3] for some recent works in this area). Micro-grids are local energy grids that can either be connected to a large distribution grid or operated autonomously, in *island mode*. They are modular systems that may include electricity generation, conversion, storage, and consumption elements. A Local Power Network (LPN) interconnects all these elements and the distribution grid. The distribution grid is the main source of electricity supply, except when the micro-grid is operating in island mode. Though micro-grids are primarily considered as electricity networks, it is worth noticing that they can include any type of local energy generation (*e.g.*, heating and cooling energy), distribution, consumption and storage elements. The micro-grid energy management problem involves two subtasks: power stabilization and cost minimization. An unbalance between supplied and consumed power tends to destabilize the power flows in the micro-grid, even leading to blackouts in extreme cases. If the micro-grid is connected to the distribution grid, the ancillary services of the distribution grid may eliminate the micro-grid unbalance automatically. The operating costs of a micro-grid can be minimized by coordinating and dispatching multiple generation, consumption and storage elements connected to the micro-grid. Main operating costs are fuel costs, costs for energy storage, and costs for electricity bought from the main distribution grid. If local generators are present in the micro-grid, the micro-grid can act as a producer and sell electrical energy to the distribution grid.

In this paper, we focus on the problem of optimal energy management for the small scale micro-grid sketched in Figure 1, which represents a building cooling system. Similar settings are considered in [4], [5], [6]. The building cooling system has no local power source and fully depends on the

main distribution grid for the electrical energy supply. Its main components are a chiller plant and a cooling load. Differently from [4], [5], [6], no storage element is present. The cooling load represents the cooling energy needed to maintain some temperature profile in a zone subject to disturbances (*i.e.*, the outside temperature and the occupancy profile), whereas the chiller plant is composed of two chillers that convert into cooling power the electrical power provided by the distribution grid through the LPN. The cooling power is then conveyed to the cooling load through the Chilled Water Circuit (CHWC).

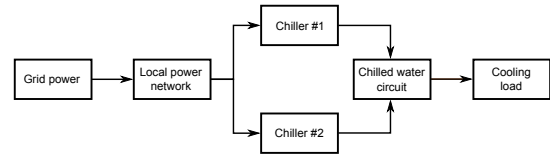


Fig. 1. Configuration of the small-scale micro-grid.

The objective is to operate the building cooling system so as to best satisfy the cooling energy demand while minimizing the electrical energy costs. This goal is pursued through two joint actions: on the one hand, the cooling power request is appropriately split between the two chillers so as to optimize the performance of the chiller plant; and, on the other hand, the zone temperature set-point is modified to some extent with respect to some reference profile so as to decrease the cooling power request and, hence, save energy. The maximal allowed variation of the set-point represents a compromise between saving and discomfort, and is the result of an agreement between the grid operator and the users. Indeed, significant energy savings can be obtained with a limited variation of the set-point, and, hence, with a limited impact on comfort. In particular, the temperature increment range is designed well within the comfort bounds set by the ISO norm on thermal comfort, [7].

The energy management problem is here formulated as a stochastic optimal control problem with constraints. A two-step solution procedure is proposed, which involves solving a nonlinear static optimization problem in the first step, and a dynamic programming problem in the second step. The latter one is tackled by Approximate Dynamic Programming (ADP) techniques (see, *e.g.*, [8], [9]) resting on the system simulation. Despite its simplicity, the considered configuration presents some of the key features that characterize a micro-grid, *i.e.*, the presence of both continuous and discrete dynamics, and of stochastic uncertainty affecting their joint evolution, which make the energy management problem challenging to solve.

Research supported by the European Commission under the MoVeS project, FP7-ICT-2009-257005.

N. Ceriani, R. Vignali, L. Piroddi, M. Prandini are with the Dipartimento di Elettronica, Informazione e Bioingegneria of the Politecnico di Milano, Via Ponzio 34/5, 20133 Milano, Italy {nceriani, vignali, piroddi, prandini}@elet.polimi.it

The rest of the paper is organized as follows. Section II provides a detailed description of the building cooling system, and the formulation of the energy management problem as an optimal control problem with constraints for a stochastic hybrid system [10]. The two-step procedure for the constrained optimal control problem is presented in Section III, whereas the ADP solution is discussed in Section IV, which includes a numerical example. Concluding remarks are given in Section V.

## II. ENERGY MANAGEMENT PROBLEM

The small-scale micro-grid under study comprises: the main distribution grid, assumed here capable of supplying an unlimited amount of power; the LPN interconnecting the main distribution grid and the power loads; the CHWC representing the cooling energy distribution system using water as the distribution medium; a chiller plant composed of two chillers that convert electrical energy into cooling energy; and the cooling load representing a zone (one or more rooms, or a partitioned space in a room) whose temperature should track a given reference profile. Cooling power is provided to this purpose by the chillers through the CHWC. The cooling power request is affected by two sources of heating power, that is the outside temperature and the internal heat gains due to the presence of people, office equipment, lighting, etc. Besides these elements, the micro-grid includes also the chilled water temperature controller and the thermostat (hereafter referred to as *lower-level controllers*), that regulate the temperatures of the water circuit and of the zone, respectively, and the energy management system to be designed, which is composed of the *temperature set-point modulator* that decides how to modify the zone temperature set-point with respect to some given reference profile so as to decrease the cooling power request, and the *chiller plant optimizer*, that decides how the cooling power request should be split between the chillers whose performance depends on the outside ambient temperature, the temperature of the cooling medium, and the requested cooling power [11].

Let  $[t_0, t_f]$  denote the reference control horizon. The zone temperature set-point modulator defines the zone temperature set-point  $T_{ZASP}$  by modifying the reference profile  $T_{ZASP}^*$  of at most some (small) amount  $\Delta_{\max}$  and only a few times during  $[t_0, t_f]$ , for a maximum total discomfort level  $d_{\max}$  representing the integral over  $[t_0, t_f]$  of the set-point increments. This translates into the following equation:

$$T_{ZASP}(t) = T_{ZASP}^*(t) + \Delta_{ZA}(t)$$

where the control variable  $\Delta_{ZA}$  of the set-point modulator is subject to the instantaneous and integral constraints:

$$0 \leq \Delta_{ZA}(t) \leq \Delta_{\max} \wedge \int_{t_0}^t \Delta_{ZA}(\tau) d\tau \leq d_{\max}, t \in [t_0, t_f].$$

A state variable  $d$  can then be introduced to account for the integral constraint on the discomfort:  $\dot{d}(t) = \Delta_{ZA}(t)$ , initialized with  $d(t_0) = 0$  and subject to  $d(t) \leq d_{\max}$ ,  $t \in [t_0, t_f]$ .

*Remark 1:* By taking  $d_{\max}$  as a measure of discomfort, it is implicitly assumed that the zone temperature behavior causes no discomfort when the zone temperature set-point is set

equal to the reference profile  $T_{ZASP}^*$  ( $d_{\max} = 0$ ) and that the zone temperature set-point is representative of the actual behavior of  $T_{ZA}$ . This entails that the lower-level control has been appropriately designed and guarantees a satisfactory tracking performance.  $\square$

Let us denote the two chillers as *Ch1* and *Ch2*. The chiller plant optimizer is assumed to instantaneously satisfy the cooling power request  $Q_{CSP}$ , compatibly with the maximum cooling power  $Q_{C,\max}$  that the chiller plant can provide:

$$Q_C(t) = \Phi_{[0, Q_{C,\max}]}(Q_{CSP}(t)), \quad (1)$$

where  $\Phi_{[a,b]}(x) = \begin{cases} a, & x < a \\ x, & x \in [a, b] \\ b, & x > b \end{cases}$  is the saturation function.

The cooling power  $Q_C(t)$  is split between the two chillers according to

$$Q_{C,Ch1}(t) = (1 - \alpha(t))Q_C(t), \quad Q_{C,Ch2}(t) = \alpha(t)Q_C(t),$$

where  $Q_{C,Ch_i}(t)$  denotes the cooling power requested to the  $i$ th chiller and  $\alpha \in [0, 1]$  is a scheduling parameter that denotes the fraction of cooling power assigned to *Ch2*, the remaining  $1 - \alpha$  fraction being requested to *Ch1*.

Before formulating the optimal energy management problem as a constrained optimization problem, we describe the equations of the controlled system –including the lower-level controllers.

### A. Controlled system equations

The electric power requested to the distribution grid is

$$P_G(t) = P_{L,Ch1}(t) + P_{L,Ch2}(t), \quad (2)$$

where  $P_{L,Ch_i}$  is the compressor power needed for chiller  $i$  to supply the cooling power  $Q_{C,Ch_i}$ .  $P_{L,Ch_i}$  is zero when chiller  $i$  is off, otherwise it depends on the outside ambient temperature  $T_{OA}$ , the CHWC temperature  $T_{CW}$ , and the evaporator duty  $Q_{C,Ch_i}(t)$  as follows:

$$P_{L,Ch_i} = \frac{a_{Ch_i,1}T_{OA}T_{CW} + a_{Ch_i,2}(T_{OA} - T_{CW})}{T_{CW} - a_{Ch_i,3}Q_{C,Ch_i}} + \frac{a_{Ch_i,4}T_{OA}Q_{C,Ch_i}}{T_{CW} - a_{Ch_i,3}Q_{C,Ch_i}} - Q_{C,Ch_i}, \quad (3)$$

where  $a_{Ch_i,k}$ ,  $k = 1, \dots, 4$ , are suitable parameters. This equation is derived from the static nonlinear Gordon-Ng model [12].

Temperature  $T_{CW}$  evolves according to

$$C_{CW} \cdot \frac{dT_{CW}}{dt} = Q_{ZA}(t) - Q_C(t) \quad (4)$$

where  $Q_{ZA}(t) = X_{C,Z}(t) \cdot k_{cw} \cdot (T_{ZA}(t) - T_{CW}(t))$  is the cooling power exchanged with the zone at temperature  $T_{ZA}$  and  $Q_C(t)$  is the cooling power provided by the chillers. Variable  $X_{C,Z}(t)$  represents the valve opening ( $X_{C,Z} = 0$  if the valve is closed and 1 if it is completely open). Heat losses in the circuit are neglected.  $T_{CW}$  is kept at some set-point value  $T_{CWSP}$  through the joint action of a proportional controller and a load disturbance compensator:

$$Q_{CSP}(t) = k_{P,CW}(T_{CW}(t) - T_{CWSP}(t)) + Q_{ZA}(t),$$

where  $k_{P,CW}$  is the proportional gain. Thus, if the chiller plant can satisfy the cooling power request (i.e.  $Q_C = Q_{CSP}(t)$ ), in view of equation (4), one has:

$$C_{CW} \frac{dT_{CW}}{dt} = -k_{P,CW} (T_{CW}(t) - T_{CWSP}(t)).$$

The zone temperature  $T_{ZA}$  evolves according to

$$C_{ZA} \frac{dT_{ZA}}{dt} = -Q_{ZA}(t) + Q_{OA}(t) + Q_{INT}(t),$$

where  $Q_{OA}(t) = k_{out}(T_{OA}(t) - T_{ZA}(t))$  is the power exchanged with the outside ambient, whereas  $Q_{INT}(t)$  is the internal heat gain contribution, described here by means of an empirical model:

$$Q_{INT}(t) = (p_1 T_{ZA}^2(t) + p_2 T_{ZA}(t) + p_3) n_{\text{people}}(t),$$

where  $n_{\text{people}}(t)$  is the zone population. Appropriate values for coefficients  $p_1$ ,  $p_2$  and  $p_3$  are suggested in [13]. The valve opening  $X_{C,Z}$  is regulated by the cooling thermostat, implementing the following PI controller with anti-wind up to account for the actuator constraint  $X_{C,Z} \in [0, 1]$ :

$$\dot{z}_{C,Z}(t) = -\frac{k_{I,Z}}{k_{P,Z}} \cdot z_{C,Z}(t) + \frac{k_{I,Z}}{k_{P,Z}} \cdot X_{C,Z}(t)$$

$$X_{C,Z}(t) = \Phi_{[0,1]}(k_{P,Z} \cdot e_{C,Z}(t) + z_{C,Z}(t))$$

$$e_{C,Z}(t) = T_{ZA}(t) - T_{ZASP}(t),$$

where  $z_{C,Z}(t)$  is the PI state variable, and  $k_{P,Z}$  and  $k_{I,Z}$  are the PI gains.

As for the disturbances  $T_{OA}$  and  $n_{\text{people}}$ , some temperature forecast is used to represent  $T_{OA}$ , whereas  $n_{\text{people}}(t)$  is described as a birth-death process with time-varying birth and death rates  $\lambda_{IN}(t)$  and  $\lambda_{OUT}(t)$ , modeling the arrival and departure of people. Rates are designed so that the resulting average occupancy matches some reference profile. The overall controlled system is a stochastic hybrid system with three continuous state variables ( $T_{CW}$ ,  $T_{ZA}$ , and  $z_{C,Z}$ ) and one discrete state variable ( $n_{\text{people}}$ ). As for the on/off status of the chillers, it is modeled implicitly through the scheduling variable  $\alpha$  (when  $\alpha = 0$ , then, *Ch1* is on and *Ch2* is off, and viceversa when  $\alpha = 1$ ). The stochastic inputs acting on the system are given by the Poisson processes determining the evolution of the birth-death process  $n_{\text{people}}$  modeling the number of occupants.

### B. Stochastic optimal control problem with constraints

Let  $s$  denote the state of the system, as defined in the previous subsection, plus the discomfort state variable  $d$ , and  $\mathcal{S}$  the state space. Our goal is to design a (deterministic) state-feedback control policy  $\pi : \mathcal{S} \times [t_0, t_f] \rightarrow [0, 1] \times [0, \Delta_{\max}]$  that maps  $s \in \mathcal{S}$  and  $t \in [t_0, t_f]$  into an appropriate value for the scheduling parameter  $\alpha \in [0, 1]$  and the set-point control variable  $\Delta_{ZA} \in [0, \Delta_{\max}]$  to be applied at time  $t$  when the state value is equal to  $s$ . The objective is to minimize the energy cost spent over the time horizon  $[t_0, t_f]$ , while not exceeding the maximum discomfort level  $d_{\max}$ , which can be expressed as the following constrained optimization

problem:

$$\min_{\pi} E_{s_0}^{\pi} \left[ \int_{t_0}^{t_f} c_G(t) P_G(t) dt \right] \quad (5)$$

subject to:  $d(t) \leq d_{\max}, \forall t \in [t_0, t_f]$ ,

where  $P_G(t)$  denotes the power requested to the main distribution grid and  $c_G(t)$  the price per unitary power request, at time  $t \in [t_0, t_f]$ .

Here,  $s_0$  denotes the state value at time  $t_0$  and  $E_{s_0}^{\pi}$  denotes the expected value when the initial state is  $s_0$  and the control policy  $\pi$  is applied. If the energy price  $c_G(t)$  is taken to be constant, one is actually minimizing the average electrical energy consumption.

### III. A TWO-STEP SOLUTION BASED ON NONLINEAR STATIC OPTIMIZATION AND DYNAMIC PROGRAMMING

The problem of designing a control policy  $\pi : \mathcal{S} \times [t_0, t_f] \rightarrow [0, 1] \times [0, \Delta_{\max}]$  solving the constrained optimization problem (5) can be decomposed into two phases without affecting the optimality of the solution:

1. design of the map  $\pi_{\alpha} : \mathcal{S} \times [t_0, t_f] \rightarrow [0, 1]$  for the scheduling of the chillers;
2. design of the map  $\pi_{\Delta_{ZA}} : \mathcal{S} \times [t_0, t_f] \rightarrow [0, \Delta_{\max}]$  for discomfort modulation.

The policy  $\pi$  is then obtained by combining these two maps:  $\pi = (\pi_{\alpha}, \pi_{\Delta_{ZA}}) : \mathcal{S} \times [t_0, t_f] \rightarrow [0, 1] \times [0, \Delta_{\max}]$ .

Indeed, observe that by equations (2) and (3),  $P_G$  is a static function of  $\alpha$ ,  $Q_C$ ,  $T_{OA}$  and  $T_{CW}$ :

$$P_G = P_G(\alpha, Q_C, T_{OA}, T_{CW}).$$

Now, since  $Q_C$ ,  $T_{OA}$  and  $T_{CW}$  are independent of  $\alpha$ , one can design the optimal scheduling strategy  $\alpha^*(Q_C, T_{OA}, T_{CW}) = \arg \min_{\alpha} P_G$ . This amounts to solving a (nonlinear) static optimization problem. The resulting  $\alpha^*(Q_C, T_{OA}, T_{CW})$  can be rewritten as a time-varying function of the state  $s \in \mathcal{S}$  since  $T_{OA}$  is a given signal, whereas  $T_{CW}$  is a state variable and  $Q_C$  is a static function of the state, thus defining the optimal map  $\pi_{\alpha}^*$ .

The map  $\pi_{\Delta_{ZA}} : \mathcal{S} \times [t_0, t_f] \rightarrow [0, \Delta_{ZA}]$  for the set-point modulation can then be designed by solving the constrained optimization problem:

$$\min_{\pi_{\Delta_{ZA}}} E_{s_0}^{\pi_{\Delta_{ZA}}} \left[ \int_0^{t_f} c_G(t) P_G^*(t) dt \right] \quad (6)$$

subject to:  $d(t) \leq d_{\max}, \forall t \in [t_0, t_f]$ ,

where  $P_G^* = P_G(\alpha^*, Q_C, T_{OA}, T_{CW})$  is the power demand when the optimal scheduling policy  $\pi_{\alpha}^*$  obtained in phase 1 is used.

Under the assumptions that: *i*) the set-point is not modulated continuously but is changed every  $T = \frac{t_f - t_0}{N}$  minutes, and *ii*) the control variable  $\Delta_{ZA}$  can take only a finite set of values, then, problem (6) can be rephrased as a finite-horizon control problem for a discrete time stochastic system with a discrete control input set. The discrete time stochastic system executions are obtained by sampling the executions of the original continuous time system, with the understanding that the control input is held constant over each  $[t_0 + kT, t_0 + (k+1)T)$  time frame.

Let  $u_k := \Delta_{ZA}(t_0 + kT)$  and  $x_k := s(t_0 + kT)$  respectively denote the control and state variables at the discrete time instant  $k$ ,  $\mathcal{U}$  the discrete control space, and  $w_k$  the stochastic inputs affecting the system evolution in the  $k$ -th time frame  $[t_0 + kT, t_0 + (k+1)T)$ . Then, the sampled version  $z_k := d(t_0 + kT)$  of the discomfort variable evolves according to  $z_{k+1} = z_k + u_k T$ , starting from  $z_0 = 0$ . Based on the introduced notation, problem (6) can be rewritten as follows:

$$\min_{\nu} E_{s_0}^{\nu} \left[ \sum_{k=0}^{N-1} c_k(x_k, u_k, w_k) \right] \quad (7)$$

subject to:  $z_k \leq d_{\max}, \forall k \in [0, N]$ ,

where  $\nu = (\nu_0, \nu_1, \dots, \nu_{N-1})$  with  $\nu_k(\cdot) = \pi_{\Delta_{ZA}}(\cdot, kT) : \mathcal{S} \rightarrow \mathcal{U}$  denotes the discrete time policy corresponding to the piecewise constant continuous time policy  $\pi_{\Delta_{ZA}}$ , and

$$c_k(x, u, w) = \int_{t_0+kT}^{t_0+(k+1)T} c_G(t) P_G^*(t) dt$$

is the one-step-cost representing the energy cost spent in  $[t_0 + kT, t_0 + (k+1)T)$  when  $s(t_0 + kT) = x$ ,  $\Delta_{ZA}(t_0 + kT) = u$ , and the stochastic inputs within  $[t_0 + kT, t_0 + (k+1)T)$  are given by  $w$ .

Problem (7) can be solved, in principle, through the Dynamic Programming (DP) approach. Here, we shall refer to the so-called  $Q$ -iteration method.

Let us define the  $Q$ -functions  $Q_k : \mathcal{S} \times \mathcal{U} \rightarrow \mathbb{R}_+$ ,  $k = 0, 1, \dots, N-1$ , where  $Q_k(x, u)$  represents the expected cost incurred over the time window  $[t_0 + kT, t_f)$ , when the state at time  $t_0 + kT$  is  $x$ , the control input at time  $t_0 + kT$  is set equal to  $u$  and maintained constant for the next  $T$  minutes, and an optimal modulation policy is applied from time  $t_0 + (k+1)T$  onwards. The  $Q$ -functions can be computed according to the following backward iterative procedure:

$$Q_k(x, u) = E_{w_k} \left[ c_k(x, u, w_k) + \min_{u' \in \mathcal{U}(z')} Q_{k+1}((\tilde{x}', z'), u') \right],$$

for  $k = 0, 1, \dots, N-2$ , initialized at  $k = N-1$  with

$$Q_{N-1}(x, u) = E_{w_{N-1}} [c_{N-1}(x, u, w_{N-1})].$$

In the former equation,  $x \in \mathcal{S}$  is the state at time  $t_0 + kT$  and  $x' = (\tilde{x}', z') \in \mathcal{S}$  denotes the value taken by the next state (*i.e.*, the state at time  $t_0 + (k+1)T$ ) when the control input at time  $t_0 + kT$  is set equal to  $u$  for the next  $T$  minutes,  $z'$  being the value of the discomfort variable.  $E_{w_k}[\cdot]$  denotes the expectation with respect to the stochastic inputs that are responsible for the probabilistic evolution of the system in the  $k$ -th time-frame  $[t_0 + kT, t_0 + (k+1)T)$ . As for  $\mathcal{U}(z')$ , it represents the set of admissible values for the control input  $u$  when the value of the discomfort variable is  $z'$  and is introduced to account for the constraint  $z_k \leq d_{\max}, \forall k \in [0, N]$ , in the DP equations. For instance, if  $\mathcal{U} = \{0, \frac{\Delta_{\max}}{2}, \Delta_{\max}\}$ , then,

$$\mathcal{U}(z') = \begin{cases} \mathcal{U}, & d_{\max} - z' \geq \Delta_{\max} T \\ \{0, \frac{\Delta_{\max}}{2}\}, & \frac{\Delta_{\max}}{2} T \leq d_{\max} - z' < \Delta_{\max} T \\ \{0\}, & d_{\max} - z' < \frac{\Delta_{\max}}{2} T. \end{cases}$$

Based on the  $Q$ -functions, the optimal policy  $\nu_k^*$  for set-point modulation can be expressed as

$$\nu_k^*(x) \in \arg \min_{u \in \mathcal{U}(z)} Q_k(x, u), \quad x = (\tilde{x}, z) \in \mathcal{S}.$$

Unfortunately, determining an exact analytical solution to the DP equations is impracticable due to the following two main reasons: *i*) the  $Q$ -iteration involves computing the expected value  $E_{w_k}[\cdot]$  with respect to the stochastic inputs affecting the system dynamics; and *ii*) given that the state space has a continuous component, numerical computations require the  $Q$ -function to be finitely parameterized.

#### IV. ADP SOLUTION

##### A. Certainly equivalence-based approach

The certainly equivalence approach to the DP solution consists in neglecting the uncertainty affecting the system evolution by replacing the stochastic inputs  $w_k$  with their nominal (deterministic) component  $\bar{w}_k$ , and computing the optimal policy  $\bar{\nu}_k : \mathcal{X} \rightarrow \mathcal{U}$ ,  $k = 0, 1, \dots, N-1$ , for the obtained deterministic system. This way, the computation of the expected values in the DP equations is avoided and the  $Q$ -iteration algorithm is reformulated as follows:

$$\begin{aligned} \bar{Q}_k(x, u) &= c_k(x, u, \bar{w}_k) + \min_{u' \in \mathcal{U}(z')} \bar{Q}_{k+1}((\tilde{x}', z'), u'), \\ k &= 0, 1, \dots, N-2, \\ \bar{Q}_{N-1}(x, u) &= c_{N-1}(x, u, \bar{w}_{N-1}), \end{aligned} \quad (8)$$

where  $x' = (\tilde{x}', z')$  is the state at the discrete time instant  $k+1$  when inputs  $u$  and  $\bar{w}_k$  are applied from  $x$  at the discrete time instant  $k$ . Based on the so-obtained  $\bar{Q}$ -functions, we can compute the discomfort modulation policy:

$$\bar{\nu}_k^*(x) \in \arg \min_{u \in \mathcal{U}(z)} \bar{Q}_k(x, u), \quad x = (\tilde{x}, z) \in \mathcal{S}, \quad (9)$$

which is optimal for the nominal system, but sub-optimal for the original stochastic system.

Note that in order to solve equations (8) numerically, we shall partition the continuous state space and take a grid point for each element of the partition. The recursion is then computed over the chosen grid points, taking the computed  $\bar{Q}$ -function at the earlier step as constant over each element of the partition. Each iteration involves computing the one-step cost, which is done by simulating the system over the corresponding time-frame.

The actual performance of the policy can be evaluated by estimating the average cost  $E_{s_0}^{\bar{\nu}} [\sum_{k=0}^{N-1} c_k(x_k, u_k, w_k)]$  through Monte Carlo simulations. This involves running  $n$  simulations of the system fed by  $n$  independently extracted realizations of the stochastic input  $w_k^{(i)}$ ,  $i = 1, 2, \dots, n$ , and computing the empirical mean  $\frac{1}{n} \sum_{i=1}^n (\sum_{k=0}^{N-1} c_k(x_k^{(i)}, \bar{\nu}_k(x_k^{(i)}), w_k^{(i)}))$ , where  $x^{(i)}$  is the state realization associated with realization  $w^{(i)}$  when policy  $\bar{\nu}$  is applied.

##### B. Numerical results

We next show the results obtained in a specific instance of the considered scenario. The zone is occupied only during the day, and the chiller plant is active from 6 a.m. to 10 p.m.

The cost of electrical power is assumed constant and unitary for simplicity. The reference zone temperature profile is constant and equal to  $T_{ZA\ SP}^* = 20\text{ }^\circ\text{C}$ . The modulation policy can modify the zone temperature set-point every  $T = 30\text{ min}$ , with a maximum instantaneous increase of  $\Delta_{\max} = 1\text{ }^\circ\text{C}$  and  $\mathcal{U} = \{0, 0.5, 1\}$ . The maximum discomfort level is  $d_{\max} = 6\text{ }^\circ\text{C}h$  corresponding to an increase of  $1\text{ }^\circ\text{C}$  for 6 hours. The nominal behavior of the disturbance inputs ( $T_{OA}$  and  $n_{\text{people}}$ ) is included in Figure 6, whereas Figure 2 plots some occupancy profiles. A list of the system parameter values is given below:

|                    |              |  |
|--------------------|--------------|--|
| Zone               | $C_{ZA}$     | $6092 \cdot 10^3\text{ J}^\circ\text{C}^{-1}$  |
|                    | $k_{out}$    | $462.5\text{ W}^\circ\text{C}^{-1}$            |
| Thermostat         | $k_{P,Z}$    | $1\text{ }^\circ\text{C}^{-1}$                 |
|                    | $k_{I,Z}$    | $0.5\text{ }^\circ\text{C}^{-1}\text{ s}^{-1}$ |
| CHWC               | $C_{CW}$     | $1.31 \cdot 10^6\text{ J}^\circ\text{C}^{-1}$  |
|                    | $k_{cw}$     | $3.29 \cdot 10^3\text{ W}^\circ\text{C}^{-1}$  |
| CHWC controller    | $k_{P,CW}$   | $86 \cdot 10^3\text{ W}^\circ\text{C}^{-1}$    |
|                    | $T_{CW\ SP}$ | $10\text{ }^\circ\text{C}$                     |
| Chiller 1          | $a_{Ch1,1}$  | $0.0056\text{ W}^\circ\text{C}^{-1}$           |
|                    | $a_{Ch1,2}$  | $10.11\text{ W}$                               |
|                    | $a_{Ch1,3}$  | $7\text{ }^\circ\text{C W}^{-1}$               |
|                    | $a_{Ch1,4}$  | $0.9327$                                       |
| Chiller 2          | $a_{Ch2,1}$  | $0.0109\text{ W}^\circ\text{C}^{-1}$           |
|                    | $a_{Ch2,2}$  | $20.22\text{ W}$                               |
|                    | $a_{Ch2,3}$  | $3.807\text{ }^\circ\text{C W}^{-1}$           |
|                    | $a_{Ch2,4}$  | $0.9325$                                       |
| Internal heat gain | $p_1$        | $-0.2199\text{ W}^\circ\text{C}^{-2}$          |
|                    | $p_2$        | $5.0597\text{ W}^\circ\text{C}^{-1}$           |
|                    | $p_3$        | $84.9168\text{ W}$                             |

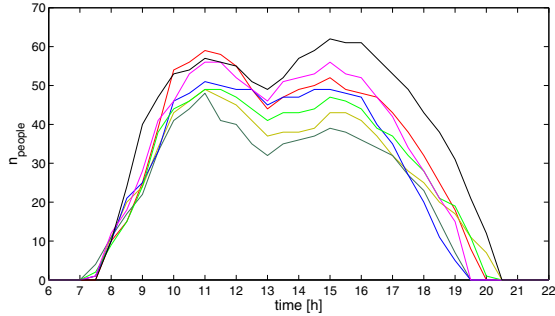


Fig. 2. Realizations of the birth-death process modeling the occupancy.

The Coefficient Of Performance,  $COP = \frac{Q_C \cdot Ch}{P_G \cdot Ch}$ , of the two chillers employed in this study is shown in Figure 3. Notice that the chillers are characterized by a different performance and chiller 1 has a higher COP for lower values of the requested cooling power.

Figure 4 represents the optimal value for  $\alpha$  as a function of  $Q_C$  and  $T_{OA}$  when  $T_{CW} = 10^\circ\text{C}$ . Correspondingly, The optimal COP for the chiller plant is given by  $COP = \frac{Q_C}{P_G^*}$ , where  $P_G^*$  is the electric power demand when the optimal scheduling policy is used. A 3D plot of the COP as a function of  $Q_C$  and  $T_{OA}$  when  $T_{CW} = 10^\circ\text{C}$  is drawn in Figure 5. We have applied the ADP approach by solving equations (8) numerically with a gridding of the continuous state variables. More in detail, the state vector is  $(T_{ZA}, X_{C,Z}, T_{CW}, d)$  where  $T_{ZA}$  is gridded on  $[19.5, 21.5]$  with a step of  $0.5^\circ\text{C}$ ,  $X_{C,Z}$  on  $[0, 1]$  with a step of 0.05. As for  $T_{CW}$  only the grid point  $T_{CW\ SP}$  is employed. Finally  $d$  is already discrete.

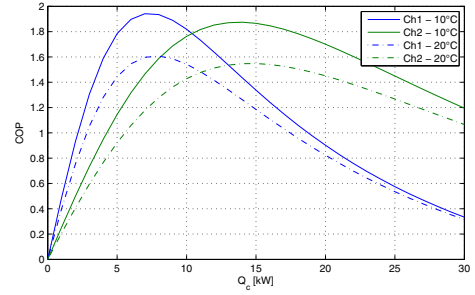


Fig. 3. COP profiles of the two chillers when  $T_{CW} = 10^\circ\text{C}$ .

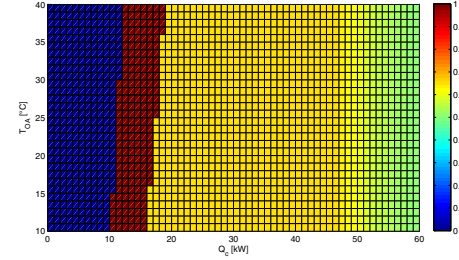


Fig. 4. Optimal value for  $\alpha$  as a function of  $Q_C$  [kW] and  $T_{OA}$  [ $^\circ\text{C}$ ] when  $T_{CW} = 10^\circ\text{C}$ .

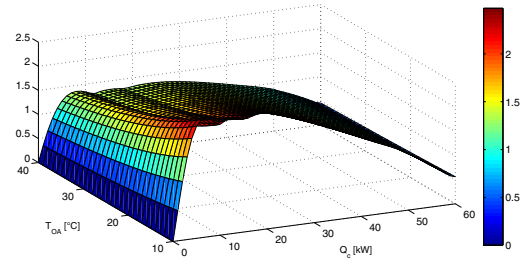


Fig. 5. Optimal COP of the chiller plant as a function of  $Q_C$  [kW] and  $T_{OA}$  [ $^\circ\text{C}$ ] when  $T_{CW} = 10^\circ\text{C}$ .

The resulting optimal policy has been tested on the system. Results obtained in the nominal case are reported in Figure 6, where the zone temperature behavior is plotted together with its set-point, the nominal occupancy and outside temperature profiles. Figure 6 shows that the policy sets an increment to the zone temperature set-point from 10:30 onwards, with a maximal increment of  $1^\circ\text{C}$  with respect to the reference set point  $T_{ZA\ SP}^* = 20\text{ }^\circ\text{C}$  in the last time slot at 21:30, when the zone is empty, and the outside ambient temperature is low. This very last increment of  $1^\circ\text{C}$  can be justified by noticing that whenever the set-point is increased, then, the chiller plant is not required to provide any cooling power until the set-point is reached. This is also the case for the adopted  $0.5^\circ\text{C}$  increment at 10:30, but at 21:30 the power saving is higher since it takes more time to the zone temperature to reach the set-point given that disturbances (outside temperature and zone occupancy) are smaller. A similar situation in terms of occupancy and temperature can be observed early in the morning. However, a set-point increase at that early stage would not be as convenient

since, due to the  $d_{\max}$  constraint, it should be followed by a set-point decrease, which would cause some cooling power consumption. Also, note that in the time intervals from 8 to 10 and from 16 to 19 the occupancy is quite similar, but the set-point is increased only in the latter time interval. This is probably due to the fact that from 8 to 10 the outside ambient temperature is smaller and, hence, the chiller plant can operate more efficiently since its COP is higher.

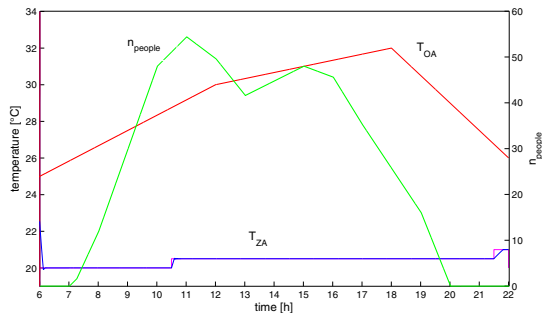


Fig. 6. Zone temperature behavior when the ADP policy is applied to the system operating in nominal conditions.

We have estimated the average energy consumption over  $10^4$  occupancy profiles extracted independently and we found it equal to  $75.28 \text{ kWh}$ . The modulation of the zone temperature set point yields an energy saving of about 3.29% with respect to a policy in which the temperature set point is maintained constant to  $T_{ZA}^* = 20 \text{ }^\circ\text{C}$ , the average energy consumption in this latter case being  $77.84 \text{ kWh}$ . Further improvements are possible if the constraints on the discomfort are relaxed.

In order to evaluate the performance gains achieved with the ADP-based control policy, the latter has been compared with non-optimized policies where the scheduling of the chillers is defined based on heuristics and the zone temperature set-point is not modulated. In particular, the following 2 scheduling strategies have been considered:

- S1) the cooling power is shared equally between the chillers;
- S2)  $Ch1$  is used up to  $10 \text{ kW}$  request,  $Ch2$  is used if  $10 \text{ kW} < Q_C \leq 30 \text{ kW}$ ; finally, if  $Q_C > 30 \text{ kW}$ ,  $Ch2$  operates at full range and  $Ch1$  supplies the cooling power request exceeding  $30 \text{ kW}$ .

Note that the scheduling strategy S2 partially exploits the knowledge on the COP profiles of Figure 3. The first naive strategy results in an average energy consumption of  $132.60 \text{ kWh}$ . This figure is significantly larger than the energy consumption  $77.84 \text{ kWh}$  relative to the optimal scheduling strategy adopted in the ADP-policy, which in fact achieves a saving of 41.29%. Strategy S2 instead provides a similar consumption to the optimal one. This is due to the fact that the two scheduling strategies operate almost identically for low cooling loads. For larger loads, such as for a 4-times larger occupancy profile than the one in Figure 6, the ADP-based scheduling policy outperforms S2 ( $159.6 \text{ kWh}$  against  $178 \text{ kWh}$  of overall energy consumption) as shown in Figure 7.

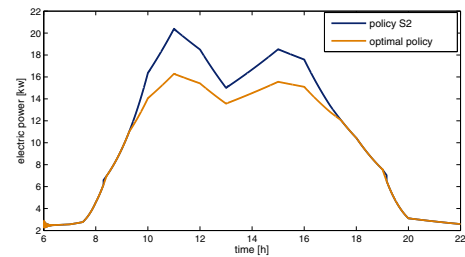


Fig. 7. Electric power request for a large cooling load.

## V. CONCLUSIONS

In this paper, we considered the optimal energy management of a building cooling system, formulating it as a stochastic optimal control problem with constraints. A solution procedure was suggested based on nonlinear static optimization and ADP techniques. The proposed framework can be extended to the case of a larger system that includes further components, such as thermal storage, and local electrical power generation and consumption. The presence of additional elements offers additional flexibility to the system, but also makes the energy management problem more challenging due to the higher dimensionality of the system. This is subject of ongoing investigations.

## ACKNOWLEDGMENTS

The authors thank Martin Strelec (Honeywell) for his support on the case study formulation and modeling.

## REFERENCES

- [1] P. Stluka, D. Godbole, and T. Samad, "Energy management for buildings and microgrids," in *Joint IEEE CDC and ECC*, 2011.
- [2] A. Parisio and L. Glielmo, "Energy efficient microgrid management using model predictive control," in *Joint IEEE CDC and ECC*, 2011.
- [3] R. Minciardi and R. Sacile, "Optimal control in a cooperative network of smart power grids," *IEEE Systems Journal*, vol. 6, no. 1, pp. 126–133, 2011.
- [4] Y. Ma, F. Borrelli, B. Hency, A. Packard, and S. Bortoff, "Model predictive control of thermal energy storage in building cooling systems," in *Joint IEEE CDC and ECC*, 2009.
- [5] Y. Ma, F. Borrelli, B. Hency, B. Coffey, S. Bengea, and P. Haves, "Model predictive control for the operation of building cooling systems," in *American Control Conference*, 2010.
- [6] M. Behl, T. Nghiem, and R. Mangharam, "Green scheduling for energy-efficient operation of multiple chiller plants," in *33rd IEEE RTSS*, 2012.
- [7] ISO Norm 7730, *Ergonomics of the thermal environment – Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria*. Geneva, Switzerland: ISO, 2005.
- [8] L. Busoniu, R. Babuska, B. De Schutter, and D. Ernst, *Reinforcement Learning and Dynamic Programming Using Function Approximators*. CRC Press, 2010.
- [9] W. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. John Wiley & Sons, 2007.
- [10] J. Lygeros and M. Prandini, "Stochastic hybrid systems: a powerful framework for complex, large scale applications," *European Journal of Control*, vol. 16, no. 6, pp. 583–594, 2010.
- [11] J. Gordon, K. Ng, and H. Chua, "Optimizing chiller operation based on finite-time thermodynamics: universal modeling and experimental confirmation," *International Journal of Refrigeration*, vol. 20, no. 3, pp. 191–200, 1997.
- [12] P. Sreedharan and P. Haves, "Comparison of chiller models for use in model-based fault detection," in *Int. Conf. for Enhanced Building Operations*, 2001.
- [13] *CIBSE Guide A: Environmental Design*. Norwich, UK: CIBSE Publications, 2006.