2018 European Control Conference (ECC)
June 12-15, 2018. Limassol, Cyprus

A Heuristic approach for Online Distributed Optimization of Multi-Agent Networks of Smart Sockets and Thermostatically Controlled Loads based on Dynamic Average Consensus

Mauro Franceschelli, Alessandro Pilloni, Andrea Gasparri

Abstract— **This paper presents a novel heuristic online optimization method and multi-agent control architecture to optimize the Peak-to-Average power Ratio (PAR) of a large population of Thermostatically Controlled Loads (TCLs) over a sliding receding horizon time window. The proposed architecture exploits only local measurements of the TCL power consumption with no knowledge of their internal temperature. No centralized aggregator of information is used and agents preserve their privacy by cooperating only through consensus-based distributed estimation. TCLs interactions occur via Smart Power Sockets (SPSs) which are interconnected through a peer-to-peer (P2P) network over the internet. The control architecture is designed from a multi-agent perspective in which real household appliances can interact with each other via SPSs.**

Our contribution is twofold: first we introduce a novel hybrid modelling of the TCL-plus-SPS system along with a method for parameter identification and a method estimate the internal state of the TLC through SPS performed power measurements; then we provide a heuristic algorithm for online distributed optimization of the on/off states of the SPSs which exploits a dynamic average consensus algorithm to estimate the planned future average power consumption of the network while preserving the agents' privacy. Numerical simulations and preliminary experimental results performed in a novel low cost testbed are provided.

I. Introduction

Matching power generation and consumption is the fundamental problem of the power grid [1], this issue is widely known to be increased by volatile renewable power generation and hourly variations of urban power demand due to time-correlation of the usage of domestic electric appliances for water heating, air conditioning, cooking etc.. This problem may be ameliorated by increasing the flexibility of power demand consumption to reduce short term electric load variations.

Thermostatically Controlled Loads (TCLs) such as water heaters, freezers, boilers and electric radiators are characterized by a simple ON/OFF power consumption dynamics and are widely accepted as devices whose active control can provide regulation capability and ancillary services to the smart grid, see [2], [3], [4], [5]. TCLs can act as energy storage devices by modulating their ON/OFF time intervals due to their slow internal temperature dynamics, thus large populations of actively controlled TCLs can effectively add some degree of freedom to the modulation of the urban electric power demand.

Since domestic TCLs consume almost half of urban electric power demand in residential areas (especially where the infrastructure for delivering natural gas to households is not developed) they have attracted significant attention of the research community in the attempt to smooth out the load variations, reduce the Peak to Average power Ratio (PAR) [6], [7] of the daily electric consumption profile and thus increase the predictability of short term power consumption. Some promising methods for electric Demand Side Management (DSM) that focus on controlling TCLs can be found in [8], [9], [4]. There, some kind of centralized strategies and distributed decision making methods supported by a centralized information aggregator are exploited to address the problem.

In this paper we are interested in a method to optimize online the behavior of a large population of TCLs to ameliorate urban electric load variations by reducing their PAR within a receding horizon time window. In particular, we aim to reduce the maximum planned power consumption in each time window while enforcing constraints on the temperature ranges of the TCLs, thus not altering the average power consumption in each interval.

To achieve this, we propose to interconnect a large population of TCLs to a P2P network over the internet via cheap Smart Power Socket (SPS) systems and then model the network as a multi-agent system where decision making is local and information available to the agents is limited to a few neighbors, to preserve user privacy and avoid the need of a centralized information aggregator managed by a third party. Therefore, by constraining ourselves to a multi-agent control architecture we avoid the cost of centralized monitoring of the network by a third party which seeks to profit from the control of the network.

The proposed approach exploits a dynamic average consensus algorithm to allow the estimation of the average planned power consumption of the whole network of TCLs by exploiting only local information on the power consumption plan of each TCL, known by each agent. Then, each agent may update in real-time its own short term consumption plan to reduce the peaks of the estimated future power consumption of the network. It follows that the dynamic average consensus algorithm enables the dynamic estimation of future power consumption concurrently with

The research leading to these results has received funding from the Italian Ministry of Research and Education (MIUR), under call "*Scientific Independence of young Researchers*" SIR, 2014, project *CoNetDomeSys*, code RBSI14OF6H.

Mauro Franceschelli and Alessandro Pilloni are with the Department of Electrical and Electronic Engineering, University of Cagliari, 09123, Italy. Email: {mauro.franceschelli, alessandro.pilloni,pisano}@diee.unica.it.

Andrea Gasparri is with the Department of Engineering, Roma Tre University, Rome, 00146, Italy. Email: gasparri@dia.uniroma3.it.

respect to each agent attempt to optimize its own planning. The reader is referred to [10], [11], [12], [13], [14] for a comprehensive overview on dynamic average consensus.

In this work we adopt the so-called multi-stage dynamic average consensus algorithm proposed in [14] due to its robustness to agents joining and leaving the network, the ability to tune the algorithm to trade-off the amount of local communication among agents for steady state estimation error, robustness to re-initialization errors, and to communication failures/changes in the network topology. These are some of the problems found in real scenario were actual devices operate in a real P2P over internet network.

A further issue that we aim to address with our control architecture arises from the use of SPSs to control TCLs. SPSs are a cheap way to retrofit existing appliances and make them "smart", thus significantly reducing the cost barrier to the introduction of wide spread electric DSM programs whose effectiveness is only as good as high is the number of users involved, thus the need to avoid expensive custom made appliances or costly infrastructure.

On the other hand, a SPS is only able to measure the power consumption of the TCL which it is connected to, not the internal temperature or state of the TCL. For this reason, in this paper we also propose a identification method for the dynamics of the TCL, modelling and a local observer design to allow real-time state estimation of each TCL via power consumption measurements only. Here, we extend the identification method proposed in [15] to the case were neither the internal temperature, nor the desired maximum and minimum temperature range of the TCL are available/known. Finally, we describe our experimental testbed and show some preliminary experimental results.

To summarize, the **main contributions** of this paper are:

- A novel heuristic method for real-time PAR optimization of large populations of TCLs modelled as a multi-agent system;
- A multi-agent control architecture which address the needs of user privacy and avoids centralized monitoring or information aggregation by third parties;
- A plug-and-play architecture to retrofit existing TCLs by the addition of SPSs to reduce the hardware requirements and thus facilitating the introduction of wide spread electric DSM strategies by the users;
- Methods for parameter identification, modeling and observer-based state estimation for the hybrid system composed by a TCL and a SPS;
- A description of our novel testbed together with some preliminary experimental results.

The reminder of the paper is organized as follows. In Section II we discuss the modelling of each agent as a TCL-plus-SPS hybrid system. In Section III the problem statement is outlined. Section IV presents the multi-agent control architecture under study. In Section V the novel heuristic algorithm based on dynamic average consensus is presented and discussed. In Section VI numerical simulations and the low cost testbed are presented, corroborated by some preliminary experimental results. Finally, concluding remarks are given in Section VII.

II. Modelling of Thermostatically Controlled Loads and Smart Power Sockets

Let us consider a Multi-Agent System (MAS) consisting of a population $\boldsymbol{\mathcal{V}} = \{1, \dots, n\}$ of TCL-plus-SPS systems. SPS are provided with a power sensor, computation and WiFi communication capabilities, and on/off switch for actuation. Thus, the SPSs may enable communication among agents within a P2P network over the internet. We denote with $\boldsymbol{\mathcal{E}} \subseteq \{\boldsymbol{\mathcal{V}} \times \boldsymbol{\mathcal{V}}\}$ the set of active communication links at time t. The network topology is modeled by a graph $\boldsymbol{\mathcal{G}} = (\boldsymbol{\mathcal{V}}, \boldsymbol{\mathcal{E}})$. According to related works [2], [16], the temperature variations of the i-th TCL can be well approximated at discrete time instants $t_1, t_2, \dots t_k$, $k \in \mathbb{N}$ as

$$\begin{array}{rcl} T^i(t_{k+1}) & = & T^i(t_k) \cdot e^{-\alpha^i dt} + \\ & & \left(1 - e^{-\alpha^i \Delta t}\right)\left(T^i_\infty + \frac{q^i}{\alpha^i} u^i(t_k)\right), \end{array} \tag{1}$$

where $dt = t_{k+1} - t_k$ is a constant sampling interval, $T^i(t_k)$ is the internal temperature of the i-th TCL at time t_k, $u^i(t_k) : \mathbb{R}^+ \to \{0, 1\}$ is the $\{\mathrm{OFF}, \mathrm{ON}\}$ state of the electric heater of the TCL. Since a SPS is connected to the power outlet of the TCL it follows that

$$u^i(t_k) = s^i(t_k) \cdot \delta^i(t_k) \tag{2}$$

where $s^i(t_k) : \mathbb{R}^+ \to \{0, 1\}$ denotes the on/off boolean state of the SPS and $\delta^i(t_k) : \mathbb{R}^+ \to \{0, 1\}$ represents the on/off state of the thermostatic heating element of the appliance. Clearly, the heating actuation can be ON only if the SPS is ON, and the electric heater is ON as well. T^i_∞ is the ambient temperature of the i-th TCL, $\alpha_i \in \mathbb{R}^+$ is the i-th heat exchange coefficient with the eternal environment, q^i is the heat generated by the electric heater and $p^i \in \mathbb{R}^+$ is the power consumption of the TCL when turned ON. TCLs are characterized by simple hysteretic control logics used to keep the internal temperature within a desired range, i.e. $T_i \in \left[T^i_{\min}, T^i_{\max}\right]$, such that

$$\begin{cases} \text{if } T^i(t_k) > T^i_{\max} & \Rightarrow \text{then } \delta^i(t_{k+1}) = 0 \\ \text{if } T^i(t_k) < T^i_{\min} & \Rightarrow \text{then } \delta^i(t_{k+1}) = 1 \\ \text{if } T^i(t_k) \in \left[T^i_{\min}, T^i_{\max}\right] & \Rightarrow \text{then } \delta^i(t_{k+1}) = \delta^i(t_k) \end{cases} \tag{3}$$

When $u^i(t_k) = 1$ (ON) the power consumption of the TCL is $p^i(t_k) = \mathrm{p}^i$ watts, when $u^i(t_k) = 0$ (OFF), $p^i(t_k) = 0$.

***Remark** 2.1:* In our application scenario temperatures $T^i(t_k)$, dead-band constraints $\left[T^i_{\min}, T^i_{\max}\right]$, ambient temperatures T^i_∞, the heat exchange coefficient α^i and the heat generated by the TCLs actuators q^i are not known a-priory and can not be measured directly. In Section IV-A an approach to estimate and identify these information only through the power consumption measurements sensed by the SPS devices is discussed in detail. ■

III. Problem Statement

Let $\tau(t_k) = [t_k, t_k + Ldt)$ be an interval of time with length Ldt equal to a receding horizon time window.

Our objective is to design a local interaction protocol among SPSs connected to a P2P network to enable an optimized planning of their on/off state $s^i(t_k)$ over a sliding time window $\tau(t_k)$ to reduce the predicted peaks of planned power consumption of the network of TCLs, subject to their local temperature constraints (1)-(3) which impose a given average power consumption over large intervals of time, thus effectively reducing the short term PAR of the networked system.

Let us now introduce some preliminary information to outline the optimization problem introduced above.

In accordance with the models described in (1)-(3), the overall instantaneous power consumption of a population $\boldsymbol{\mathcal{V}}$ of TCLs equipped with SPSs at time t_k is given by

$$P(t_k) = \sum_{i \in \boldsymbol{\mathcal{V}}} p^i(t_k) = \sum_{i \in \boldsymbol{\mathcal{V}}} \mathrm{p}^i \cdot u^i(t_k). \tag{4}$$

Let us denote the predicted power consumption of the network at a future time $t = t_k + \ell dt$ as

$$P_\ell(t_k) = \sum_{i \in \boldsymbol{\mathcal{V}}} p^i u^i_\ell(t_k) = \sum_{i \in \boldsymbol{\mathcal{V}}} \mathrm{p}^i \cdot s^i_\ell(t_k)\delta^i_\ell(t_k). \tag{5}$$

At each time instant t_k, the generic agent i actuates the status of its SPS according to the value of $s^i_1(t_{k-1})$ and computes a new planning $s^i_1(t_k), s^i_\ell(t_k), \ldots, s^i_L(t_k)$. To simplify the notation, we use vector notation for the planning variables as follows $\mathbf{s}^i = \left[s^i_1; \ldots, s^i_\ell; \ldots; s^i_L\right]$ and omit the dependence of the planning vector from time. Vector $\mathbf{s}^i \in \{0, 1\}^L$ represents the on/off planned state of the i-th SPS in the receding horizon time window $\tau(t_k)$. At each instant t_k, every agent i aims to optimize its own scheduling of operations $\mathbf{s}^i$ by interacting with other agents and optimizing the next objective function, representing the future peak power consumption of the network within the time window $\tau(t_k)$, as consequence of their emergent behavior

$$\mathbf{s}^i(t_{k+1}) = \underset{\mathbf{s}^i \in \chi^i(t_k);\, i \in \boldsymbol{\mathcal{V}}}{\operatorname{argmin}} \ \max_{\ell = 1, \ldots, L} \ P_\ell(t_k) \tag{6}$$

where P_ℓ is defined in (5) and $\chi^i(t_k)$ is a time dependent set of linear constraints of the i-th TCL that depend upon its characteristic parameters as described in Section IV-B, the current estimated state of the TCL and the auxiliary variables u^i, δ^i needed to model its temperature dynamics in the receding horizon time window.

At each instant t_k, we enable each SPS to compute an approximate solution $\boldsymbol{s}^{i,\star}(t_k)$ of problem (6), then each SPS updates its state according to $s^i(t_{k+1}) = s^{i,\star}_1(t_k)$.

Notice that $\chi^i(t_k)$ is a set of linear constraints with mixed real, and boolean variables, (1)-(3), thus problem (6) is NP-hard in general.

IV. Multi-agent Control Architecture

In this section we describe the proposed multi-agent control architecture enabled by smart power sockets on existing domestic TCLs, such as electric water heaters. SPSs add communication, processing and measurement capabilities to existing appliances, thus enabling feedback and cooperative control within a P2P network over the internet. In the following we show how each agent, in the proposed architecture, models the combination of a TCL and a SPS, and how the parameter identification, state observation and mixed logic dynamical modeling problems can be addressed.

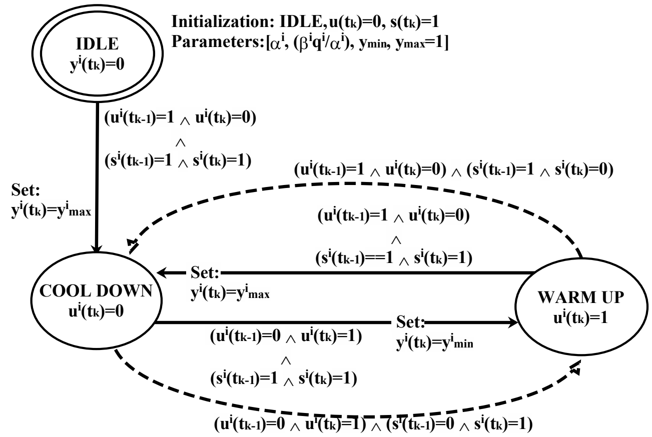

Fig. 1. Local model-based hybrid virtual temperature observer.

A. System Identification and Temperature Estimation

TCL systems control the temperature of a fluid by a hysteretic control logic, i.e. a thermostat, as in (3). To predict their power consumption, it is required to predict when they are switched ON and OFF. This can be achieved by estimating their internal temperature and identifying their characteristic parameters of their dynamics shown in (1)-(3).

1) TCL Parameters Identification: To identify the characteristic parameters of the TCL dynamics from power consumption measurements only, we measure the charge/discharge intervals of time of (1)-(3).

First, we consider a TCL with internal temperature equal to the ambient temperature, this is achieved by simply switching off the SPS for a long time. It follows that $T^i(0) = T^i_\infty$. The TCL then switches on, and the SPS measures the absorbed power up to $T^i(\Delta t^i_1) = T^i_{max}$, then the TCL switch off again. Δt^i_1 denotes the *cold charge time* and it can be evaluated by the SPS. Then, the SPS measures the time it takes for the device to switch on again and denotes this *discharge time* as Δt^i_2. Follows that $T^i(\Delta t^i_1 + \Delta t^i_2) = T^i_{\min}$. Finally, we measure the time it takes for the TCL to charge up from $T^i_{\min}$ to $T^i_{\max}$. We denote this *charge time* as Δt^i_3. By substituting in (1), it can be shown that the next relationship between the considered intervals of time holds

$$e^{-\alpha^i \Delta t^i_2} = \frac{T^i_{\max} - T^i_\infty}{T^i_{\min} - T^i_\infty} = \frac{\cancel{\frac{q^i}{\alpha^i}}\left(1 - e^{-\alpha^i\left(\Delta t^i_1 - \Delta t^i_3\right)}\right)}{\cancel{\frac{q^i}{\alpha^i}}\left(1 - e^{-\alpha^i \Delta t^i_1}\right)}. \tag{7}$$

The above equation can then be solved numerically to compute the only unknown variable α^i.

Now, the temperature range $\left[T^i_{\min}, T^i_{\max}\right]$ is unknown and it is a private information of the user which can not be estimated by the SPS via power consumption measurement only. On the other hand our model does not need to estimate the absolute temperature of the device by merely the instants of time in which its heating element switches on and off. Therefore, we build an equivalent model of the dynamics of a virtual temperature which varies in a normalized interval with the purpose to predict the future instants of time in which the TCL is "on". Let $y^i(t_k) = \beta^i(T^i(t_k) - T^i_\infty)$ be a coordinate change representing a virtual internal temperature such that $y^i : \left[T^i_\infty, T^i_{\max}\right] \mapsto \left[y^i_\infty = 0, y^i_{\max} = 1\right]$. By substitution into (1), it yields

$$y^i(t_{k+1}) = y^i(t_k)e^{-\alpha^i dt} + \frac{\beta^i q^i}{\alpha^i}\left(1 - e^{-\alpha^i dt}\right)u^i(t_k) \quad (8)$$

Now we consider the discharge time Δt^i_2. By eq. (8) with $y^i = 1$ and $u^i = 0$, it holds

$$y^i_{\min} = e^{-\alpha^i \Delta t^i_2}.$$

Finally, by considering the charge time Δt^i_1 from $y^i_{\min} = e^{-\alpha^i \Delta t^i_2}$ to $y^i_{\max} = 1$, by manipulating (8) it holds

$$\frac{\beta^i q^i}{\alpha^i} = \frac{1 - e^{-\alpha^i \Delta t^i_2} e^{-\alpha^i \Delta t^i_3}}{\left(1 - e^{-\alpha^i \Delta t^i_3}\right)} \quad (9)$$

Thus, by following the proposed procedure, all parameters of the virtual temperature model of the TCL in (8) can be identified by power consumption measurement only. The virtual temperature model enables a model-based prediction of future "on/off" switching times of the TCL.

2) Model-based Hybrid Virtual Temperature Observer: We now consider the problem of estimating the virtual temperature $y^i(t_k)$ of the model given in eq. (8) by measuring only power consumption of the TCL via the SPS. The state variables of interest to the observer are the state of the socket $s^i(t_k)$ and the ON/OFF state $\delta^i(t_k)$ of the TCL, available by measuring power consumption. Clearly, the TCL can be ON only if the SPS is ON while the TCL is OFF necessarily if the SPS is OFF. In other words, the actual state of the TCL can be seen as the product of two boolean variables $u^i(t_k) = s^i(t_k)\delta^i(t_k)$ as in (2).

The hybrid system representing the observer is shown in Figure 1. The observer exploits information about the thermostatic control method of the TCL described in (3). It consists of three discrete states and a continuous state $y^i(t_k)$ which corresponds to the virtual temperature of the TCL according to (8).

In the "IDLE" state the observer detects the falling edge from $u^i(t_{k-1}) = 1$ to $u^i(t_k) = 0$ when the SPS is ON $s^i(t_{k-1}) = s^i(t_k) = 1$, we call this a "synchronization" event. When this event is detected, it means that the TCL is OFF because $T^i(t_k) = T_{\max}$. Thus, the observer sets the virtual temperature $y^i(t_k)$ to $y^i_{\max} = 1$ and initiate the observation process by moving to the "COOL DOWN" state, which represents the temperature discharge process given by (8), with $u^i = 0$. The "WARM UP" state models the charging process of y^i. If a "synchronization" event is detected, the temperature is set to $T^i(t_k) = T_{\max}$ and the state moves to "COOL DOWN". In the "COOL DOWN" state and "WARM UP" states the virtual temperature evolves in open loop as in (8) until another synchronization event is detected. A typical behavior of the observed temperature is shown in Figure 1.

***Remark** 4.1:* Notice that the proposed observer may stay in "IDLE" for long time before it detects the trigger condition that sets the estimated temperature to the correct value, i.e., $y^i = 1$. Clearly, only agents for which the observers are synchronized with their TCLs can be involved in the optimization task. This is not an issue for the proposed architecture since it is based on a plug-and-play approach where agents may join and leave with no need of algorithm re-initialization. it is worth to remark that the observer estimates $y^i(t_k)$ according to an open-loop model (as per eq. (8)) in between synchronization events, thus estimation errors due to parameter uncertainties are unavoidable. On the other hand, for the correct functioning of the method we only need a short term prediction of the TCLs behavior within a time-frame up to at most a couple of hours, i.e., a single discharge cycle. Experimental results demonstrate that these approximations are sufficient to ensure cooperation among TCLs. Finally, notice that although we assumed in (8) the sampling frequency $dt = t_{k+1} - t_k$ constant and synchronous for each TCL, this restriction can be easily removed in practice by simply measuring the time interval past among two subsequent records. Future work will focus on increasing the accuracy of the identification procedure and to generalize the results to the asynchronous case.

B. Mixed Logic Dynamical modeling of TCL-plus-SPS

In this section we use Mixed Logic Dynamical System (MLD) modeling [17] to define a set of boolean variables and linear constraints, in particular the set $\chi^i(t_k)$ in problem (6), that models the behavior of the TCL plus SPS hybrid system. The MLD modeling tool allows us to translate state transitions in a hybrid system expressed as boolean functions into linear integer constraints with boolean variables. Due to space limitations we do not report here all the manipulations needed to compute a MLD model and limit ourselves to show the final result.

Next we show the set of inequalities needed to represent the local dynamical constraints of the TCL-plus-SPS hybrid system, according to [17], based on the hysteretic control law (3) and the SPS state, represented by the boolean variables $u^i(t_k)$, $s^i(t_k)$, $\delta^i(t_k)$ and the continuous state variable $y^i(t_k)$ of each agent

$$\begin{aligned} -s^i(t_k) + u^i(t_k) &\le 0 \\ -\delta^i(t_k) + u^i(t_k) &\le 0 \\ s^i(t_k) + \delta^i(t_k) - u^i(t_k) &\le 1 \end{aligned} \quad (10)$$

$$\begin{aligned} -g^i(t_k) + \epsilon + \delta^i_1(t_k)(g^i_{\min} - \epsilon) &\le 0 \\ -f^i(t_k) + f^i_{\min}(1 - \delta^i(t-1)) + h^i_{\min}\,\delta^i_2(t_k) &\le 0 \\ \delta^i_1(t_k) + \delta^i_2(t_k) - 2\,\delta(t_k)^i &= 0 \end{aligned} \quad (11)$$

where $g^i_{\min}$, $f^i_{\min}$, and $h^i_{\min}$, are abbreviations for the minimum of, respectively, functions $g^i(t_k) = (y^i(t_k) - y_{\min})$, $f^i(t_k) = (y^i(t_k) - 1)$, and $h^i(t_k) = f^i(t_k) - f^i_{\min}(1 - \delta^i(t_{k-1}))$. In addition, $\epsilon \in \mathbb{R}^+$ is an arbitrarily small constant, whereas δ_1^i and δ_2^i denote dummy variables required to express the inequalities as linear constraints.

In addition, the set of local dynamical constraints $\chi^i(t_k)$ in (6) includes also the constraints due to the linear dynamics of the normalized temperature in (8) and the bounds of its range of variation $y^i \in [y^i_{\min}, 1]$. These constraints are provided as inequality constraints on the control input, i.e., $\mathbf{A}^i\mathbf{u}^i \leq \mathbf{b}^i$. From (8), to simplify the notation we let $\hat{A}^i = e^{-\alpha^i dt}$ and $\hat{B}^i = \frac{\beta^i q^i}{\alpha^i}(1 - e^{-\alpha^i dt})$. Then, by denoting $\mathbf{u}^i$, and $\mathbf{y}^i$, and $\boldsymbol{\delta}^i$ as vectors of L entries with corresponding elements u^i_ℓ, y^i_ℓ and δ^i_ℓ, $\ell = 1, \ldots, L$, we can predict the trajectory of $y^i(t_k)$ over the receding horizon time window $\tau_k(t_k) = [t_k, t_k + Ldt)$, as follows

$$\mathbf{y}^i = \underbrace{\begin{bmatrix} \hat{B}^i & 0 & \cdots & 0 \\ \hat{A}^i\hat{B}^i & \hat{B}^i & \cdots & 0 \\ \vdots & & \ddots & \\ \hat{A}^{iL-1}\hat{B}^i & \hat{A}^{iL-1}\hat{B}^i & \cdots & \hat{B}^i \end{bmatrix}}_{\boldsymbol{F}^i} \mathbf{u}^i + \underbrace{\begin{bmatrix} \hat{A}^i \\ \hat{A}^{i2} \\ \vdots \\ \hat{A}^{iL} \end{bmatrix}}_{\boldsymbol{G}^i} y^i(t_k). \quad (12)$$

The local temperature constraints can be compactly represented as $\mathbf{A}^i\mathbf{u}^i \leq \mathbf{b}^i$ where matrix $\mathbf{A}_i$ and vector $\mathbf{b}_i$ are

$$\mathbf{A}_i = \begin{bmatrix} \boldsymbol{F}^i \\ -\boldsymbol{F}^i \end{bmatrix}, \quad \mathbf{b}_i = \begin{bmatrix} y^i_{\max}\mathbf{1} - \boldsymbol{G}^i y^i(t_k) \\ -y^i_{\min}\mathbf{1} + \boldsymbol{G}^i y^i(t_k) \end{bmatrix}. \quad (13)$$

Finally, the set of linear constraints $\chi^i(t_k)$ is properly composed to include all the previously mentioned logical and dynamical constraints in, resp., (10), (11), and (12), (13).

V. Proposed Algorithm

In this section we present the proposed heuristic approach to approximate a solution to problem (6) with our multi-agent control architecture.

Algorithm 1, named "TCL Cooperation Protocol", consists of a local state update rule executed by each agent indefinitely. Each agent owns a local prediction of the future average power consumption of the network over the horizon $\tau(t_k)$. To update this prediction each agent gathers periodically the prediction made by neighboring agents and then updates its own by exploiting the multi-stage dynamic consensus algorithm proposed in [14]. At each iteration, each agent attempts to minimize, with probability μ, the maximum predicted power consumption of the network over $\tau(t_k)$, by updating its own scheduling of on/off states according to the local constraints (10)-(13), thus solving problem (6). Notice that, the local constraints of each agent $\chi^i(t_k)$ are time-varying since they depend upon the current state of the TCL at the time the optimization takes place.

Note that, although the optimization problem in (14) is in general NP-hard, as it involves mixed linear and boolean variables, in our setting for each agent i, the number of variables to be optimized is relatively small as only local constraints over a short time horizon $\tau(t_k)$ are involved. For example, about 20/60 steps into the future may account for 30 minutes to two-three hours of operations, depending on the tuning of the algorithm. Thus, the complexity of the proposed method does not increase with the size of the network, but only with respect to the size of the time horizon.

Furthermore, it should be noticed that the solution does not need to be optimal, i.e., since the approach is iterative and ultimately heuristic. As a matter of fact, what we need is a feasible solution which improves the objective function value with respect to the current scheduling of the SPS. To achieve this, we exploit a standard branch and bound solver with a fixed time limit. In section VI we provide a detailed discussion of the numerical simulations and preliminary experiments carried out along with the tuning parameters that

Algorithm 1 TCL Cooperation Protocol

Estimation variables of Agent i:
$P^i_{\ell,w}(k)$, for $w = 1, \ldots, m$, $\ell = 1, \ldots, L$;
$\mathbf{s}^i = [s^i_1, \ldots, s^i_\ell, \ldots, s^i_L]$;
ON/OFF Scheduling of Agent i:
$u^i_\ell(t_k)$, $s^i_\ell(t_k)$, $\ell = 1, \ldots, L$;
Tuning parameters:
$\rho, \epsilon \in \mathbb{R}^+$, $\epsilon < \frac{1}{2D_{\max}}$, $\rho < 1 - \epsilon D_{\max}$, $\mu, \xi > 0$;
$D_{\max}$: maximum degree of $\boldsymbol{\mathcal{G}}$;
Initial counters: $h = 0$, $k = 0$;
Protocol execution:
All agents repeat indefinitely the next operations, here reported for agent i:
• Increment the iteration counter $h = h + 1$;
• Gather $P^j_{\ell,w}(h)$ for $\ell = 1, \ldots, L$, $w = 1, \ldots, m$, from current neighbors $j \in \boldsymbol{\mathcal{N}}_i$;
• Update state variables $P^i_{\ell,w}(h)$ for $\ell = 1, \ldots, L$, $w = 1, \ldots, m$ as follows

$$P^i_{\ell,1}(h+1) = P^i_{\ell,1}(h) - \sum_{j \in \mathcal{N}_i} \varepsilon \left(P^i_{\ell,1}(h) - P^j_{\ell,1}(h)\right) + \rho\left(p_i u^i_\ell(t_k) - P^i_{\ell,1}(h)\right)$$
$$P^i_{\ell,2}(h+1) = P^i_{\ell,2}(h) - \sum_{j \in \mathcal{N}_i} \varepsilon \left(P^i_{\ell,2}(h) - P^j_{\ell,2}(h)\right) + \rho\left(P^i_{\ell,1}(h) - P^i_{\ell,2}(h)\right)$$
$$\vdots$$
$$P^i_{\ell,m}(h+1) = P^i_{\ell,m}(h) - \sum_{j \in \mathcal{N}_i} \varepsilon \left(P^i_{\ell,m}(h) - P^j_{\ell,m}(h)\right) + \rho\left(P^i_{\ell,m-1}(h) - P^i_{\ell,m}(h)\right)$$

• With probability μ, solve the next optimization problem

$$[\mathbf{s}^{i,\star}, \mathbf{u}^{i,\star}] = \underset{\mathbf{s}^i \in \chi(t_k)}{\operatorname{argmin}} \max_{\ell \in \{1,\ldots,L\}} p_i u^i_\ell + P^i_{\ell,m}(h) \quad (14)$$

If $[\mathbf{s}^{i,\star}, \mathbf{u}^{i,\star}]$ is a feasible solution and

$$\max_{\ell \in \{1,\ldots,L\}} \mathrm{p}^i u^{i,\star}_\ell + P^i_{\ell,m}(h) < \max_{\ell \in \{1,\ldots,L\}} \mathrm{p}^i u^i_\ell + P^i_{\ell,m}(h) - \xi$$

then set $\mathbf{s}^i(t_k) := \mathbf{s}^{i,\star}$, $\mathbf{u}^i(t_k) := \mathbf{u}^{i,\star}$
else $\mathbf{s}^i(t_k) := \mathbf{s}^i(t_k)$, $\mathbf{u}^i(t_k) := \mathbf{u}^i(t_k)$, i.e., do nothing.
Endif
• Measure time t;
• **If** $t > t_k + dt$ **then** shift the receding horizon time window:
Set the current SPS state equal to:
$s^i(t_{k+1}) := s^i_1(t_k)$;
Shift power consumption prediction by dt:

$$\mathbf{u}^i(t_{k+1}) := \left[u^i_2(t_k), \ldots, u^i_\ell(t_k), \ldots, u^i_{L-1}(t_k), u^i_{L-1}(t_k)\right]$$

Shift SPS scheduling by dt:

$$\mathbf{s}^i(t_{k+1}) := \left[s^i_2(t_k), \ldots, s^i_\ell(t_k), \ldots, s^i_{L-1}(t_k), s^i_{L-1}(t_k)\right]$$

Let $k := k + 1$
Endif

have been chosen for our case study. Finally, if the approximate solution does not improve on the current scheduling of operation of the generic agent, the agent may simply discard it and the current scheduling is carried out until a better one is found in future iterations.

A. Convergence analysis

We now discuss the behavior of Algorithm 1. Briefly, the algorithm relies on a properly tuned dynamic average consensus algorithm to let each agent estimate with sufficient accuracy and sufficiently fast the predicted future average power consumption of the network within a given horizon.

The basic idea of the proposed algorithm is that, thanks to a properly tuned dynamic average consensus algorithm each agent can estimate with sufficient accuracy and sufficiently fast the predicted future average power consumption of the network within the time window $\tau(t_k)$. The parameters to be tuned are ρ, ε and m. These parameters are linked to the steady-state estimation error, the convergence rate and the estimation error during transient behavior, when the predicted future consumption changes due to the local optimization performed by the agents on their own schedule of operations. A discussion concerning how to tune the dynamic average consensus algorithm to achieve a given steady state error performance with guaranteed error performance during transient behavior with time-varying reference signals is given in [14]. Details concerning the tuning for this application scenario are here omitted for the sake of brevity.

Let us now state one of the main convergence properties of the proposed heuristic approach, which relies on the property of the dynamic average consensus algorithm described in [14] to guarantee a maximum estimation error ξ.

Proposition 5.1: Consider a multi-agent system executing Algorithm 1. If the estimation error at iteration h and time $t \in [t_k, t_{k+1})$ on the future average power consumption of each agent satisfies

$$\left| P^i_{\ell,m}(h) - \frac{P_\ell(t_k)}{n} \right| \leq \xi, \quad \forall \quad \ell = 1, \ldots, L$$

then the objective function of problem (6) is non-increasing inside all intervals $[t_k, t_{k+1})$ for $k = 1, \ldots, \infty$.

Proof: To prove this, it is sufficient to point out that each agent updates its own scheduling of operations only if

$$\max_{\ell \in \{1,\ldots,L\}} \mathrm{p}^i u^{i,\star}_\ell + P^i_{\ell,m}(h) < \max_{\ell \in \{1,\ldots,L\}} \mathrm{p}^i u^i_\ell + P^i_{\ell,m}(h) - \xi. \quad (15)$$

Therefore, even in the worst case of maximum estimation error, the update is executed only if

$$\max_{\ell \in \{1,\ldots,L\}} \mathrm{p}^i u^{i,\star}_\ell + \frac{P_\ell(t_k)}{n} \pm \xi - \xi < \max_{\ell \in \{1,\ldots,L\}} \mathrm{p}^i u^i_\ell + \frac{P_\ell(t_k)}{n}, \quad (16)$$

thus

$$\max_{\ell \in \{1,\ldots,L\}} \mathrm{p}^i u^{i,\star}_\ell + \frac{P_\ell(t_k)}{n} < \max_{\ell \in \{1,\ldots,L\}} \mathrm{p}^i u^i_\ell + \frac{P_\ell(t_k)}{n}. \quad (17)$$

If we now consider the maximum over the summation for all agents $i \in \mathcal{V}$ and remember that only one agent updated its scheduling, thus leaving the constant ξ outside the summation, it holds

$$\max_{\ell \in \{1,\ldots,L\}} \sum_{i \in \mathcal{V}} \mathrm{p}^i u^{i,\star}_\ell + P_\ell(t_k) < \max_{\ell \in \{1,\ldots,L\}} \sum_{i \in \mathcal{V}} \mathrm{p}^i u^i_\ell + P_\ell(t_k). \quad (18)$$

At this point, since $P_\ell(t_k) = \sum_{i \in \mathcal{V}} \mathrm{p}^i u^i_\ell(t_k)$, it holds

$$\max_{\ell \in \{1,\ldots,L\}} \sum_{i \in \mathcal{V}} \mathrm{p}^i u^{i,\star}_\ell < \max_{\ell \in \{1,\ldots,L\}} \sum_{i \in \mathcal{V}} \mathrm{p}^i u^i_\ell, \quad (19)$$

thus proving that the global objective function of problem (6) is non-increasing during the algorithm iterations and decreasing any time an agent is able to reduce its own power consumption at the peak of the planned future network power consumption, according to its own local constraints. □

Finally, it should pointed out that it holds the recursive feasibility property of the local constraints $\chi^i(t_k)$. In this preliminary paper we omit the proof and limit ourselves to point out that for a sufficiently small dt it always exist feasible scheduling of operations of the current temperature is inside the desired range. If the temperature is estimated to be outside the desired range the SPS simply stays "on" and the thermostat on the TCL behaves according to its hysteretic control law given in (3).

VI. Simulations and Experimental Testbed

In this section, we provide numerical simulations along with preliminary experimental results, carried out in a novel low-cost testbed, to corroborate the effectiveness of the proposed approach.

A. Simulations results

For the simulations, we considered a connected P2P network $\mathcal{G}(\mathcal{V}, \mathcal{E})$ of $n = 500$ agents for which the topology was generated as a random undirected Erdös-Rényi graph, with edge existence probability $3\log(n)/n$. The dynamics of the agents detailed in (1)-(3) was initialized with parameters that were chosen randomly from a set of 6 actual electric water heaters that have been used in the experiments. In particular, we chose $\mathrm{p}^i \in [1.2, 1.6]\mathrm{kW}$, $\alpha^i = 1.38 \times 10^{-4}(1 \pm 5\%)\mathrm{s}^{-1}$, $T^i_\infty = 25°\mathrm{C}$, $q^i = 0.098\ °\mathrm{C/kW}$. The initial temperature was chosen randomly within a desired temperature range $[T_{\min} = 45, T^i_{\max} = 50]°\mathrm{C}$. The parameters of the observer given in (8)-(9) were derived according to the procedure described in Section IV-A.1.

The proposed multi-agent control architecture operated in real-time with sampling-rate $dt = 3\mathrm{min}$. The receding horizon time window was chosen with length $100\mathrm{min}$, and thus it included $L = 50$ time slots. At $t = 0$, Algorithm 1, was initialized with a feasible scheduling plan to the local constraints of each agent, whereas the SPS initial plan was to stay always on, i.e., $s^i_\ell = 1$, $\ell = 1, \ldots, 50$ (see the top-plot of Figure 3 for $t = 0$). The probability of each agent i to execute the local optimization (14) at each iteration k of Algorithm 1 was set to $\mu = 0.004$.

In the top of Figure 2 it is shown an example of the actual planning at time of agent 9 at different instants of time during the algorithm execution. Furthermore, in the

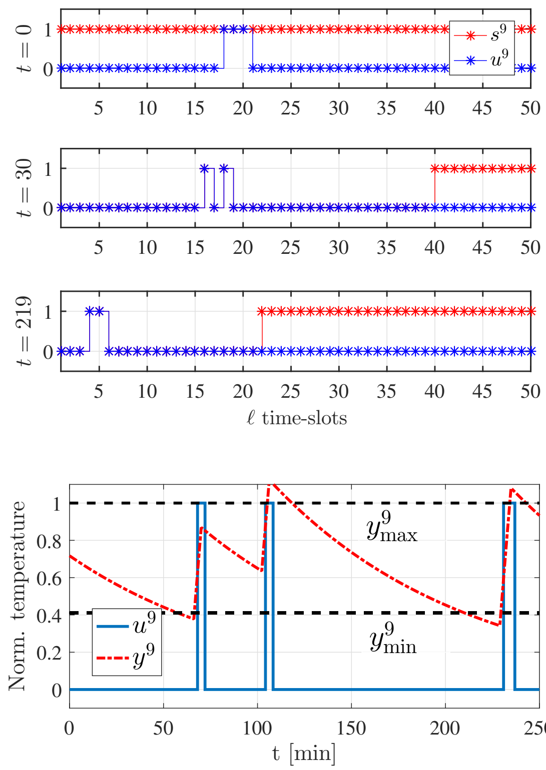

Fig. 2. Top: ON/OFF scheduling of SPS 9 over the an horizon of $L = 50$ time slots. Bottom: Estimated virtual temperature profile of TCL 9.

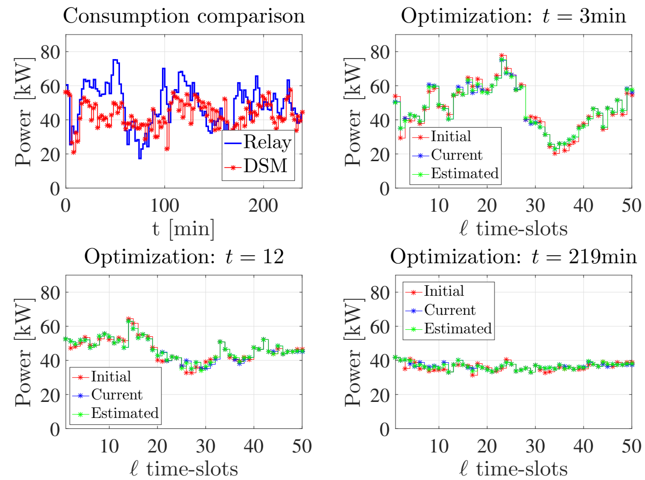

Fig. 3. Top-left: Power consumption's comparison for the network without (blue), with the proposed DSM control (red). Other plots: Profile of the initial (red), estimated (green), current (blue) objective function of problem (6) on the optimization horizon at different time $t = \{3, 12, 219\}$min.

bottom of Figure 2 it is shown the temporal profile of the virtual normalized temperature (8) of agent 9 estimated by the proposed hybrid observer, shown in Figure 1. In the top-left plot of Figure 3 it is shown a comparison between the absorbed power by the considered network of 500 TCLs coordinated by the proposed Algorithm 1, with respect to an equivalent set of autonomous TCL with traditional relay-control. In particular, it can be noticed the reduction in consumption during the whole test of 250min.

In Figure 3 results on the optimization tasks at different time interval are shown, at $t = \{3, 12, 219\}$min respectively. It can noticed that, at each step, the proposed method effectively shaves off the initial peak of the predicted power over the horizon window, depicted in red. In blue and green are shown the actual power consumption of the network over the time horizon and the predicted one computed by the dynamic average consensus consensus algorithm, respectively.

The numerical optimization involved in this simulation was solved by using the Matlab's Mixed Integer Linear Programming (*intlinprog*) solver with a time limit, i.e., a "Branch and Bound" algorithm that yields the best solution found within 5 seconds.

B. Experimental results

The preliminary results provided in this paper consider a small scale testbed composed of 6 domestic electric water heaters. Future work with existing hardware will scale the experiment up to 100 heterogenous TCLs such as water heaters and electric radiators.

Each water heater was plugged into a WeMo® Insight Smart Power Socket. The SPS has an on/off remote switch for actuation, power consumption sensing with at least 1Hz sampling frequency, and WiFi communication capabilities [18]. This off-the-shelf SPS was chosen because it offered sufficient hardware capabilities and an open API. In Figure 4 one of the experimental devices under test is shown.

The low-cost testbed used to test the proposed algorithm, named "CoNetDomeSys", has been developed in collaboration with the University of Cagliari. The testbed consists of two Java applications, a client and a server. One installed on a Raspberry "Pi Zero W" connected to the same WiFi of the smart socket, and a server application capable of managing and updating in real-time a database with all the information concerning the state of all connected smart sockets. The software has been integrated with MatLab to allow a fast prototyping and testing of the algorithms.

In Figure 5 the preliminary experimental tests are shown. The measured power consumption (blue) of a set of 6 TCLs over 150 minutes is compared with the same set of TCLs that cooperate by the proposed algorithm (red). It can be noticed how the peak power consumption of the test is significantly lowered when using the proposed algorithm compared to the case when the TCLs while the average power consumption is approximately the same.

VII. Conclusions

In this work, a novel heuristic method for real-time PAR optimization for a large populations of TCLs modelled as a MAS has been presented. A distributed agent-based architecture to address the needs of user privacy, avoid centralized monitoring while minimizing the peak of consumption has been proposed. The approach is designed for robustness with plug-and-play features to retrofit existing TCLs by

Fig. 4. Left: An electric water heater equipped with a smart power socket. Right: A Raspberry Pi Zero W and a Wemo Insight Smart.

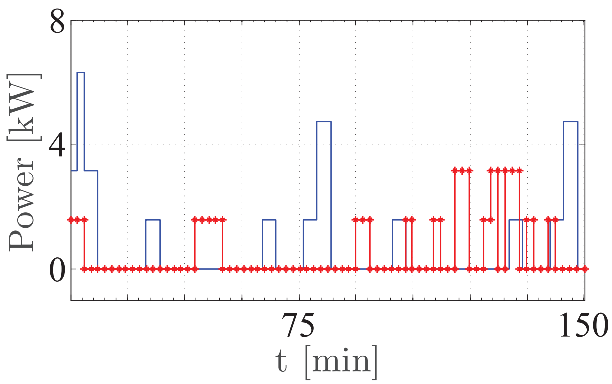

Fig. 5. Power consumption comparison between 6 autonomous TCLs (blue) and the same TCL executing the proposed algorithm (red)

the addition of low-cost smart-plugs. This framework has been developed with the intention to minimize the hardware requirements and related costs to promote the adoption of DSM strategies by the users. Main contributions of this work can be summarized as: i) an ad-hoc parameter identification method for TCLs modeled as in (8); ii) a novel model-based hybrid observer for temperature estimation which exploits only on/off power consumption information; iii) a mixed logic dynamical modelization for TCL-plus-SPS devices suitable to execute real-time MPC-like optimizations; and iv) a novel asynchronous, randomized algorithm for minimizing the peak of consumption over a prediction horizon, that exploit only local power consumption measurements as well as pair-wise-communication to perform its DSM specifications. Numerical simulation on large scale network of TCLs along with preliminary experiments carried out with our novel low-cost testbed platform are given to corroborate the effectiveness of the proposed architecture.

ACKNOWLEDGEMENTS

We thank Gianluca Mereu for his support in the management and debugging of the experimental testbed.

REFERENCES

[1] W. El-Khattam, K. Bhattacharya, Y. Hegazy, and M. Salama, "Optimal investment planning for distributed generation in a competitive electricity market," *IEEE Trans. on Power Systems*, vol. 19, no. 3, pp. 1674–1684, 2004.

[2] H. Hao, B. Sanandaji, K. Poolla, and T. Vincent, "Aggregate flexibility of thermostatically controlled loads," *IEEE Trans. on Power Systems*, vol. 30, no. 1, pp. 189–198, 2015.

[3] S. H. Tindemans, V. Trovato, and G. Strbac, "Decentralized control of thermostatic loads for flexible demand response," *IEEE Trans. on Control Systems Technology*, vol. 23, no. 5, pp. 1685–1700, 2015.

[4] S. Grammatico, B. Gentile, F. Parise, and J. Lygeros, "A mean field control approach for demand side management of large populations of thermostatically controlled loads," *IEEE European Control Conf.*, pp. 3548–3553, 2015.

[5] M. Franceschelli, A. Gasparri, and A. Pisano, "Coordination of electric thermal systems for distributed demand-side management: A gossip-based cooperative approach," 2016, pp. 623–630.

[6] A.-H. Mohsenian-Rad, V. W. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid," *IEEE Trans. on Smart Grid*, vol. 1, no. 3, pp. 320–331, 2010.

[7] I. Notarnicola, M. Franceschelli, and G. Notarstefano, "A duality-based approach for distributed min-max optimization with application to demand side management," in *IEEE Conf. on Decision and Control*, 2016, pp. 1877–1882.

[8] J. H. Braslavsky, C. Perfumo, and J. K. Ward, "Model-based feedback control of distributed air-conditioning loads for fast demand-side ancillary services," *IEEE Conf. on Decision and Control*, pp. 6274–6279, 2013.

[9] H. Xing, Y. Mou, Z. Lin, and M. Fu, "Fast distributed power regulation method via networked thermostatically controlled loads," *19th IFAC WC*, pp. 5439–5444, Aug. 2014.

[10] D. P. Spanos, R. Olfati-Saber, and R. M. Murray, "Dynamic consensus on mobile networks," 2005, pp. 1–6.

[11] M. Zhu and S. Martínez, "Discrete-time dynamic average consensus," *Automatica*, vol. 46, no. 2, pp. 322 – 329, 2010.

[12] S. S. Kia, J. Cortés, and S. Martínez, "Distributed event-triggered communication for dynamic average consensus in networked systems," *Automatica*, vol. 59, pp. 112 – 119, 2015.

[13] ——, "Dynamic average consensus under limited control authority and privacy requirements," *International Journal of Robust and Nonlinear Control*, vol. 25, no. 13, pp. 1941–1966, 2015.

[14] M. Franceschelli and A. Gasparri, "Multi-stage discrete time dynamic average consensus," *IEEE Conf. on Decision and Control*, pp. 897–903, 2016.

[15] M. Shaad, A. Momeni, C. P. Diduch, M. Kaye, and L. Chang, "Parameter identification of thermal models for domestic electric water heaters in a direct load control program," *IEEE Canadian Conf. on Electrical and Computer Eng.*, pp. 1–5, 2012.

[16] C. Perfumo, E. Kofman, J. H. Braslavsky, and J. K. Ward, "Load management: Model-based control of aggregate power for populations of thermostatically controlled loads," *Energy Conversion and Management*, vol. 55, pp. 36–48, 2012.

[17] A. Bemporad and M. Morari, "Control of systems integrating logic, dynamics, and constraints," *Automatica*, vol. 35, no. 3, pp. 407–427, 1999.

[18] Belkin, "Wemo insight smart plug." [Online]. Available: http://www.belkin.com/us/p/P-F7C029/