Control for large scale demand response of thermostatic loads

Totu, Luminita Cristiana; Leth, John-Josef; Wisniewski, Rafal

Published in:
American Control Conference (ACC), 2013

Publication date:
2013

Document Version
Accepted author manuscript, peer reviewed version

Link to publication from Aalborg University

Citation for published version (APA):
Control for large scale demand response of thermostatic loads*

Luminita C. Totu, John Leth and Rafael Wisniewski
Department of Electronic Systems, Faculty of Engineering and Science
Aalborg University
9220 Aalborg, Denmark
lct,jjl,raf@es.aau.dk

Abstract—Demand response is an important Smart Grid concept that aims at facilitating the integration of volatile energy resources into the electricity grid. This paper considers a residential demand response scenario and specifically looks into the problem of managing a large number thermostat-based appliances with on/off operation. The objective is to reduce the consumption peak of a group of loads composed of both flexible and inflexible units. The power flexible units are the thermostat-based appliances. We discuss a centralized, model predictive approach and a distributed structure with a randomized dispatch strategy.

I. INTRODUCTION

In many countries including Denmark [1], [2], energy generation from volatile resources such as wind or solar radiation is planned to increase. While these resources are sustainable and have overall capacity to cover the growing energy demand and even replace capacity currently served by fossil-fuels, large scale use is challenging. This is because the power system needs to be in balance between consumption and production at all times. When a large percentage of the generation is volatile, the balancing effort increases beyond the possibilities of the traditional grid.

Smart Grid is a developing technology that proposes real-time information exchange, distributed generation, distributed storages, and intelligent solutions for the electrical network. It can facilitate the large scale integration of volatile generation, reduce infrastructure investments and decrease the need for large, stand-by energy reserves. An important Smart Grid concept is demand response, which can be described as the active, continuous participation of the consumption in the energy balance: use less electricity when it is scarce and difficult to produce, and more otherwise.

We present a demand response scenario where a large group of thermostat-based appliances with on/off operation represent power flexible units. We can think of the flexible units as many, small and "leaky" thermal storages.

Next, we briefly refer to works on a similar topic to outline our focus. Dynamic demand, a concept closely related to demand response, is addressed in [3] for a population of domestic refrigerators acting as grid frequency stabilizers. In this case, and in contrast to the demand response scenario, the units cannot be used for planning, e.g., storing energy minutes or hours before a consumption peak.

The demand-side management structures in [4] and [5] are more appropriate for operating appliances as distributed storages. Important techniques used here are randomization and broadcasting. Furthermore, [6] discusses three different structures for demand response (price signal, individual power reference, and individual temperature reference) and concludes that a successful scheme must combine optimization and feedback. It is on these four ideas that we build the distributed approach proposed in this work.

We start by investigating centralized optimization techniques that are commonly used for production planning [7], [8]. While these can offer an insight into the consumption problem, such a direct approach alone is impractical. Due to non-convex elements (the on/off device level control) and the large number of variables that need to be communicated and computed, algorithms become impracticable. Consequently, we propose a distributed structure with two levels: a supervisor center and local controllers. The supervisor center broadcasts a global coordination signal and uses power measurements of the cumulated consumption as feedback. A modified thermostat algorithm acts as the local controller of each appliance. The algorithm handles device specific operation and responds to the coordination signal in a randomized manner.

The article is organized as follows. First, models based on physical principles are introduced in section II. The centralized optimization is presented in III, and the distributed structure in IV. Simulations for both approaches are discussed in V, while VI concludes and points to future work.

II. MODELS

We assume given \( N + M \) power consuming devices, where \( N \) units have thermal storage capabilities and thermostat driven on-off behavior, and \( M \) units have a purely stochastic, time-varying on-off behavior and no energy storage properties. We think of the first type as refrigerators, heat-pumps, air-conditioning or water boilers, and of the second type as lights, TVs, or ovens. With respect to the energy needed for nominal operation, the devices of the first category have power flexibility and are considered controllable, while those of the second category are power inflexible and are considered uncontrollable. It is also assumed that power consumption is constant during the on-cycle for all devices.

The aim is to control the \( N \) flexible devices, within the boundaries of their nominal operation and in the presence of
the $M$ inflexible devices, such that the peak of the cumulated consumption is reduced.

For both device categories, simplified physical models combined with stochastic elements are used to capture the main behaviors related to power consumption. Essential aspects of the problem are scale and variability: the objective is to manage a very large number of units and to tolerate parameter variations.

A. Basic models for the flexible units

The flexible devices have an on/off consumption pattern based on thermostat control. In the on-cycle, power is consumed and thermal energy (heat or "coldness") is stored. In the off-cycle, the thermal energy is lost in the surrounding environment. Modeling based on physical principles is described next.

The power active component (e.g. vapor-compression cycle, resistive heater, etc.) is modeled with a constant coefficient of performance. A number of compartments of uniform temperatures are modeled by heat balance affine differential equations. Since the control and communication will be based on digital systems, it is natural to work directly in discrete time. We will use models of the form (1) where the notation is summarized in Table I. A random term is introduced in the dynamics. It can be designed to account for the variety of disturbances coming from usage profiles and the environment.

\[
\begin{align*}
F_1: \quad & T_i(k+1) = A_i T_i(k) + b_i u_i(k) + c_i + q_i(k), \quad i = 1, \ldots, N \\
& y_i(k) = p_i u_i(k),
\end{align*}
\]

**Table I**

Notation and symbols for models of flexible units

<table>
<thead>
<tr>
<th>Signal</th>
<th>Description</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_i$</td>
<td>time vector for the compartments</td>
<td>$\mathbb{R}^{n_i}$</td>
</tr>
<tr>
<td>$u_i$</td>
<td>on/off value of the power consuming comp.</td>
<td>${0, 1}$</td>
</tr>
<tr>
<td>$y_i$</td>
<td>power consumption (Watts)</td>
<td>$\mathbb{R}_+$</td>
</tr>
<tr>
<td>$q_i$</td>
<td>random contributions</td>
<td>$\mathbb{R}^{n_i}$</td>
</tr>
</tbody>
</table>

**Parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_i$</td>
<td>number of flexible units</td>
<td>$\mathbb{N}$</td>
</tr>
<tr>
<td>$A_i$</td>
<td>linear map, 2D-matrix</td>
<td>$\mathbb{R}^{n_i} \rightarrow \mathbb{R}^{n_i}$</td>
</tr>
<tr>
<td>$b_i, c_i$</td>
<td>linear map, 1D-vector</td>
<td>$\mathbb{R} \rightarrow \mathbb{R}$</td>
</tr>
<tr>
<td>$p_i$</td>
<td>power rating of the device (Watts)</td>
<td>$\mathbb{R}_+$</td>
</tr>
<tr>
<td>$p$</td>
<td>collective power ratings</td>
<td>$\mathbb{R}^N$</td>
</tr>
<tr>
<td>$DT_i$</td>
<td>minimum down(off) time periods</td>
<td>$\mathbb{N}_+$</td>
</tr>
<tr>
<td>$UT_i$</td>
<td>minimum up(on) time periods</td>
<td>$\mathbb{N}_+$</td>
</tr>
</tbody>
</table>

Indexes

<table>
<thead>
<tr>
<th>Index</th>
<th>Description</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i$</td>
<td>unit index</td>
<td>$1, \ldots, N$</td>
</tr>
<tr>
<td>$k$</td>
<td>discrete time index</td>
<td>$1, 2, \ldots, K$</td>
</tr>
</tbody>
</table>

Nominal operation is the evolution of $F_1$ within a set of constraints, e.g., temperature ranges and minimum on and off times. The unit has operational flexibility because there are different possibilities of controlling the on/off power cycle, i.e. the $u_i$ signal, to maintain nominal operation.

Next, we collect the $u_i$ and $p_i$ terms in the following notations, $u(k) = (u_1(k), \ldots, u_N(k))$, $p = (p_1, \ldots, p_N)$ and write the total consumption of the $N$ flexible units at time $k$ as

\[
g(k) = \langle p, u(k) \rangle = \sum_{i=1}^{N} p_i u_i(k).
\]

B. Basic models for the inflexible units

Inflexible units have a stochastic on-off behavior with time varying properties. As example, an indoor light appliance is more likely to be turned on in the early morning and in the evening, and less at midday and after midnight. A natural choice for modeling the random on/off behavior at the unit level is using the discrete-time Markov chain formalism. We will use notations similar to [9].

Each inflexible unit $j$ will be modeled as a discrete time Markov chain, with $X_j(k) \in \{1(\text{on}), 0(\text{off})\}$ the random variable representing the state of the unit at time $k$, $P_j(k) = (p_{0j}(k), p_{1j}(k)) = (P[X_j(k) = 0], P[X_j(k) = 1])$ the state probability row vector at time $k$, and $p_{0j}(k)$ and $p_{1j}(k)$ time varying transition probabilities, "turn on" and "turn off" respectively. The evolution in time of the state probability and the power consumption output can be described as

\[
I_j: \begin{cases}
P_j(k+1) = P_j(k) M_j(k) \\
\{w_j(k) = \langle p_{0j}(k), p_{1j}(k) \rangle, 1 - p_{0j}(k)
\end{cases}
\]

where $M_j(k) = \left[ \begin{array}{cc} 1 - p_{0j}(k) & p_{0j}(k) \\ p_{1j}(k) & 1 - p_{1j}(k) \end{array} \right]$ is the transition probability matrix.

This Markov chain is also depicted in Fig. 1 and notation and symbols are summarized in Table II. The transition probabilities can be parameterized to approximate usage patterns for different device types.

**Fig. 1.** Markov chain for the inflexible units

We further collect the $X_j$ and $p_j'$ terms in the notations, $X(k) = (X_1(k), \ldots, X_M(k))$ and $p' = (p_1'(k), \ldots, p'_M(k))$ to compactly express the total power consumption of the $M$ inflexible units, a random process, as

\[
w(k) = \langle p', X(k) \rangle.
\]

In this work, we use the probabilistic construction only for numerical simulations. In the optimization formulation, a deterministic sequence $\bar{w}(k)$ is used as a forecast for the expected power consumption of the inflexible units over a required time horizon. This deterministic sequence is constructed by replacing each $w(k)$ random variable with its
### TABLE II

<table>
<thead>
<tr>
<th>Notation and Symbols for Models of Inflexible Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_j(\cdot)$</td>
<td>Probabilities for Markov states on and off</td>
</tr>
<tr>
<td>$X_j(\cdot)$</td>
<td>Random variable, state of the unit: on(1) or off(0)</td>
</tr>
<tr>
<td>$w(\cdot)$</td>
<td>Power consumption (Watts)</td>
</tr>
<tr>
<td>$\bar{w}(\cdot)$</td>
<td>Cumulated power consumption</td>
</tr>
<tr>
<td>$\mu(\cdot)$</td>
<td>Forecast/mean consumption profile (Watts)</td>
</tr>
<tr>
<td>$\mathbf{M}$</td>
<td>Number of inflexible units</td>
</tr>
<tr>
<td>$\mathbf{M}_j(\cdot)$</td>
<td>Right stochastic matrix, time varying</td>
</tr>
<tr>
<td>$p_{ij}(\cdot)$</td>
<td>Pr. of transition from state $x$ to $y$, time varying</td>
</tr>
<tr>
<td>$p_i(\cdot)$</td>
<td>Pr. of being in state $x$</td>
</tr>
<tr>
<td>$p_j(\cdot)$</td>
<td>Power rating of the device (Watts)</td>
</tr>
<tr>
<td>$\mathbf{P}$</td>
<td>Collective power ratings</td>
</tr>
</tbody>
</table>

#### Signals

- $P_j(\cdot)$: Probabilities for Markov states on and off.
- $X_j(\cdot)$: Random variable, state of the unit: on(1) or off(0).
- $w(\cdot)$: Power consumption (Watts).
- $\bar{w}(\cdot)$: Cumulated power consumption.
- $\mu(\cdot)$: Forecast/mean consumption profile (Watts).

#### Parameters

- $\mathbf{M}$: Number of inflexible units.
- $\mathbf{M}_j(\cdot)$: Right stochastic matrix, time varying.
- $p_{ij}(\cdot)$: Pr. of transition from state $x$ to $y$, time varying.
- $p_i(\cdot)$: Pr. of being in state $x$.
- $p_j(\cdot)$: Power rating of the device (Watts).
- $\mathbf{P}$: Collective power ratings.

#### Indexes

- $j$: Unit index.
- $k$: Discrete time index.

In this setup, the expected value, $E[w(k)]$, given an initial state probability vector $P_j(0)$ and knowing the transition matrices $M_j(k)$, is defined as:

$$E[w(k)] = \sum_{j=1}^{M} p_j^T P_j(0) M_j^k [0 \ 1]^T,$$

where $M_j^k = M_j(0) M_j(2) \ldots M_j(k-1)$.

Using the above models for the flexible and inflexible units, the total power consumption of the $N + M$ devices can be expressed, depending on the context, as one of the following two random processes:

- $z(k) = y(k) + w(k) = \langle p, u(k) \rangle + \langle p', X(k) \rangle$
- $\bar{z}(k) = y(k) + \bar{w}(k) = \langle p, u(k) \rangle + \bar{w}(k)$

It is noted that while $\bar{w}(k)$ has been introduced as a deterministic sequence, $\bar{z}(k)$ remains a random process for all practical cases. This is because the dynamics of the flexible units are affected by noise, and this fact will reflect into the power consumption, the term $u(k)$.

### III. A CENTRALIZED APPROACH

A straightforward approach to reduce the peak consumption is to employ optimization techniques based on an objective, models and constraints to compute the on/off controls for the flexible units. In this section, we formulate and analyze a mixed integer linear optimization that is the core component of the model predictive control (MPC) approach.

We want to calculate the $u_i(k)$ values for all the flexible units $F_i$, and over the entire time horizon $\{1, \ldots, K\}$, such that the maximum value of the sequence $\bar{z}(k)$ is minimized. This opaque minimax objective,

$$\min_{u(\cdot)} \|z(k)\|_\infty = \min_{u(\cdot)} \|\langle p, u(k) \rangle + \bar{w}(k)\|_\infty,$$

can be written in a convenient linear form [11] by adding a new variable $m \in \mathbb{R}$ and a set of inequality constraints,

$$\min_{u(\cdot)} \|z(k)\|_\infty \Rightarrow \begin{cases} \min_{u(\cdot), m} & m \in \mathbb{R} \\
- m & \leq \bar{z}(k) \leq m; \forall k. \end{cases}$$

Furthermore, because $z(k)$, the cumulated instantaneous power consumption, is always non-negative, the left inequality, $-m \leq \bar{z}(k)$, can be dismissed. Next, we detail three main constraints.

First, the model dynamics must be included to indicate the relation between the states $T_i(\cdot)$ and the decision variables $u_i(\cdot)$. For each flexible unit $i$, $i \in \{1, \ldots, K\}$, equation (1a) without the random term is accounted as an equality constraint. It is noted that the $T_i(1)$ temperatures are already decided by the $u_i(0)$, and act as initial conditions. These constraints apply on $T_i(2)$ to $T_i(K + 1)$.

Second, the states $T_i(k)$ must be within the allowable ranges, $T_i^{\min} \leq T_i(k) \leq T_i^{\max}$, where $T_i^{\min}, T_i^{\max} \in \mathbb{R}^+$, are individual parameters of the flexible unit $i$ and $k \in \{2, \ldots, K + 1\}$.

Third, we include minimum on and off times associated with the power consumption of each $F_i$. These constraints can be written in a linear form [8]. Expressions (2) are for the minimum on-time and (3) are for the minimum off-time.

$$\begin{cases} (UT_i - C_i) u_i(0) + \sum_{k=1}^{UT_i-1} u_i(k) = (UT_i - C_i) u_i(0), & (2a) \\
\sum_{k=t}^{\infty} u_i(k) \geq UT_i (u_i(t) - u_i(t-1)) & \forall t \in \{UT_i - C_i, u_i(0) + 1, \ldots, K - UT_i + 1\}, \tag{2b} \\
\sum_{k=t}^{\infty} u_i(k) - (u_i(t) - u_i(t-1)) \geq 0 & \forall t \in \{K - UT_i + 2, \ldots, K\}. \tag{2c} \end{cases}$$

It can be seen that a number of approximately $2K$ linear inequalities are used for each unit to assure the minimum on and off time conditions. The inequalities have different expressions at the beginning (2a), (3a) and at the end (2c), (3c) of the time horizon. $C_i$ is counter value that holds number of samples that the unit has been in the initial state, and is part of the optimization initialization. Thus value $(UT_i - C_i)$ represents the number of periods that the unit must remain in its initial state and not change. The term
\[ u_i(t) - u_i(t-1) \] equals to 1 for a turn-on event at time \( t \), and thus the next \( UT_i \) commands \( u_i(t+1), \ldots, u_i(t+UT_i-1) \) must remain on (=1). The term \( u_i(t-1) - u_i(t) \) equals to 1 for a turn-off event at time \( t \).

The optimization can now be passed to a mixed integer linear solver. Simulation results are presented in section V.

We present next implementation considerations. In total, the problem has \( 2 \times N \times K + 1 \) decision variables (temperatures are also decision variables) and about \( 5 \times N \times K \) constraints. A sample period \( T_s = 60 \) seconds is considered a good choice with respect to the dynamic characteristics of the controllable units, the level of detail in the modeling, and the on/off control behavior. Ideally, the time horizon \( K \) should cover 24 hours, the main period of the inflexible consumption pattern. In the model predictive control solution, such an optimization needs to be solved every sample period. Although good algorithms exist for solving mixed integer linear optimizations [12], memory and execution time requirements will grow exponentially with the number of units and the time horizon, and make computation infeasible for large scale problems. Another disadvantage here is related to the centralized nature of the approach: the computation center must send and receive data from geographically distributed locations in a short period of time, requiring a robust, fast, double way communication infrastructure.

IV. A DISTRIBUTED APPROACH

In the centralized approach, the non-convex elements and the large number of local variables and constraints make the numerical computations impracticable. It seems reasonable to carry out part of the control task at the local level where knowledge of the state variables and operational constraints is inherent. If the power consumption decisions are made locally, some type of information sharing becomes necessary to achieve a consistent global behavior.

We construct a demand response structure that has distributed characteristics, two control levels, and requires minimal communication. A diagram is shown in Fig. 2, where blocks \( F_1, \ldots, F_n \) represent flexible units equipped with the local controllers \( K_i \) and the \( (I_1 + \ldots + I_L) \) block corresponds to the cumulated inflexible consumption process.

A. Supervisor Control and Estimation

Using consumption forecasts and an aggregation model with approximate knowledge of type and number of cooling units in the system, the supervisor control generates the power reference signal \( r \). The power reference has the purpose of scheduling periods of thermal energy storage and discharge. Furthermore, estimated values are needed to characterize the refrigeration population. These are \( N^{on} \) and \( N^{off} \), estimates for the number of units in the state ON and OFF respectively that are in the flexibility range, and \( \bar{p}_i \) the average specific power rating of a cooling unit in the group. The flexibility range refers to the temperature of a unit \( i \) being some distance away from the hard limits, that is in the range \( [T_{i_{\min}} + \Delta T', T_{i_{\max}} - \Delta T'] \). This is done to avoid successive on/off cycling due to conflicting local and global objectives.

B. Local Controller

The local controller, shown in Fig. 3, is a computationally inexpensive extension of the thermostat cooling logic. The on/off decision is made with first priority on constraints, in this case temperature and timer limits. If they are all satisfied, the operational flexibility is used to respond to the broadcast signal \( \epsilon \) in a randomized manner. Fig. 4 shows a normal thermostat operation versus an extended thermostat reacting in a randomized manner to an external signal \( \epsilon \).

Fig. 2. Distributed control structure

Fig. 3. Local controller block \( K_L \)

(a) Normal duty cycle (b) Randomized duty cycle

Fig. 4. Operation of flexible unit. Temperature is shown in blue, and the on/off state in black.
C. Dispatch Strategy

The $\epsilon$ signal is similar to the error signal in a classic control structure. It is build as a scaled difference between the reference signal $r$ and the actual power consumption $z$. It can either encourage consumption ($\epsilon > 0$), or discourage it ($\epsilon < 0$). The scaling is performed such that the absolute value has the meaning of a fraction of the total number of refrigerators. For example, if $\epsilon = -0.1$, the broadcast information is that 10% of the refrigerators should turn off. The formula for computing $\epsilon$ is thus

$$
\epsilon = \begin{cases} 
(r - z)/(\tilde{N}^{\text{on}} \tilde{p}_i), & \text{if } r > z \\
(r - z)/(\tilde{N}^{\text{off}} \tilde{p}_i), & \text{if } r < z.
\end{cases}
$$

Refrigerator units in the flexibility range respond to the $\epsilon$ signal by making random trial with success rate $|\epsilon|$. If the trial is successful, the unit reacts by turning on ($r > 0$) or respectively off ($r < 0$). For a sufficiently large number of refrigerators and good $\tilde{N}$ estimates, by the law of large numbers, the cumulated responses of the individual units will be close to the requested fraction. It is thus possible to follow the a power reference signal that is well-designed.

We have organized a dispatch strategy that can support an arbitrarily large number of individual units, is computationally and communication-wise cheap, and is robust to faults in the coordination level. The dispatch is intrinsically noisy, but the relative noise ratio will decrease with the number of units. This can be seen in the next section. Some of the complexity of the problem remains to be handled at the coordination level, where good aggregation models and algorithms are needed for tracking the thermal storage level and for estimating the $\tilde{N}^{\text{on}}$ and $\tilde{N}^{\text{off}}$ values.

V. NUMERICAL EXPERIMENTS

This section puts forward specific models for the flexible and inflexible units, and presents simulation results for the centralized optimization control and the distributed control.

We take the case of a 1-compartment cooling unit in constant ambient temperature, similar to [3], [4] or [5].

$$
F_i: \begin{cases} 
T_i(k+1) = a_i T_i(k) + \\
+ (1 - a_i) (T^q_i - u_i(k) T^q_i) + q(k) \\
y_i(k) = p_i u_i(k),
\end{cases}
$$

where $a = \exp(-UA \cdot T_s/C)$, $T^q = \text{COP} \cdot p/UA$ and $T_s = 60$ seconds. A random, white component $q(k)$ with normal distribution has been added to the temperature dynamics to simulate disturbances. Overall, this is a simple model with the purpose of evaluating the control approaches, and is a particular case of the general affine linear state-space representation (1).

We work with two parameter sets, described in Table III. Parameters marked with * will be generated with a $\pm 10\%$ normal variation around the given value. In uncontrolled operation mode, the thermostat drives the power consumption cycle in regular intervals of approximately 20 minutes on and 162 minutes off for the first parameter set (refrigerators) and

<table>
<thead>
<tr>
<th>C*</th>
<th>heat capacity (^{\circ}\text{C} \cdot \text{m}^-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UA*</td>
<td>overall heat transfer (^{\circ}\text{C} \cdot \text{m}^-1)</td>
</tr>
<tr>
<td>T*\text{min}</td>
<td>ambient temp. (^{\circ}\text{C})</td>
</tr>
<tr>
<td>COP*</td>
<td>coefficient of perf.</td>
</tr>
<tr>
<td>$p^*$</td>
<td>power rating W</td>
</tr>
<tr>
<td>T*$\text{min}$</td>
<td>min operational temp. (^{\circ}\text{C})</td>
</tr>
<tr>
<td>T*$\text{max}$</td>
<td>max operational temp. (^{\circ}\text{C})</td>
</tr>
<tr>
<td>UT*</td>
<td>min up time</td>
</tr>
<tr>
<td>DT*</td>
<td>min down time</td>
</tr>
<tr>
<td>$q(\cdot)$</td>
<td>random, normal distrib.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Units</th>
<th>Refriger.</th>
<th>Freezer</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.4 $10^4$</td>
<td>1.432</td>
<td>0.8</td>
</tr>
<tr>
<td>24</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>-2.8</td>
<td>2.8</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>140</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-22</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-15</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

43 minutes on and 490 minutes off for the second parameter set (freezers).

The inflexible unit $j$ is described by the transition probabilities $p_{ij}(k)$ and $p_{j0}(k)$ specified over the time horizon of a day. The transition probabilities have been designed piece-wise constant, and create a state probability profile $p_j(k)$ with two exaggerated peaks, shown in Fig. 5. The probability behavior is identical for all units, such that the cumulated consumption has the same shape.

For the MPC approach, the optimization problem defined in Section III has been implemented for the Gurobi solver MATLAB interface. It was parameterized for a group of identical 1-compartment cooling units with identical ambient temperatures and inflexible consumption profile $\tilde{w}(k)$. The largest number of controllable units for which it was possible to obtain solutions in a reasonable amount of time was $N = 20$. The optimized commands associated corresponding to the first time step, $u(1)$, are dispatched to each unit. The new temperatures are then collected and used as initial conditions for the next optimization, which is carried out with a receded horizon.

For the distributed approach, a simulation was set-up in MATLAB as described in section IV. The elements of the Supervisor Control and Estimation block have not been completely developed for this simulation. The reference signal $r$ was designed manually such that periods of energy storage were scheduled prior to the peaks, and periods of energy discharge were scheduled during the peaks. In addition, the $N^{\text{on}}$ and $N^{\text{off}}$ estimates were constructed using simulation data that would not normally be available to the supervisor center.

![Fig. 5. Properties for the inflexible units. In the left figure, Turn On event probabilities are shown in red and Turn Off event probabilities in blue.](image-url)
Fig. 6 shows realizations for the small scale case. The differences between the planned schedule Fig. 6(a) and the MPC realization 6(b) are due to model parameter variations and noisy dynamics in the simulation of the flexible units, and also because of errors in the inflexible consumption forecast. For the distributed control 6(d), it can be seen that the reference signal for the total consumption is followed in an imprecise manner. This is mainly due to large relative granularity of the system as the response from the refrigerator group is in steps of approximately \(\pm 100\text{W}\). Furthermore, at this scale, the dispatch strategy is imprecise due to the randomization. For both the MPC and the distributed control, the peak has been reduced.

![Fig. 6. Small scale experiments for refrigerator units, \(N = 20\). Blue shows the consumption of the flexible units \(x(t)\), red shows the consumption of the inflexible units \(w(t)\), and black represents the total power \(z(t)\). For the distributed control, the power reference \(r(t)\) is shown in green. In figures 6(c) and 6(d), data from 3 days is plotted in an overlapped manner.](image)

Fig. 7 shows realizations for the large scale case. The dispatch strategy is now able to keep the response smooth and precise. The shortcomings in this experiment are only related design of the power reference signal. This remains to be addressed in future work. It can also be noticed that the second peak starting around 16:40 has a longer duration. The refrigerator units simply do not have enough storage capacity to completely ride through this peak and after the 20:00 mark, power consumption increases over the reference value. Fig. 8 shows a distributed control run for a group of freezers, units with a larger individual flexibility. In this case, it is possible to keep the peak close to the level of the inflexible consumption.

![Fig. 8. Distributed control for freezer units, \(N = 1000\).](image)

### VI. CONCLUSION

We have proposed in this work a demand response scenario with a power peak reduction objective and emphasis on large scale. First, a centralized MPC framework was evaluated as infeasible. Subsequently, we presented a distributed control structure with good performance for large numbers of control units. A main contribution is the randomized dispatch strategy, which keeps decision making at the local level where specific operation is easy to manage. The approach has only minimal communication requirements. Future work will address in detail the design of the supervisor level control based on an aggregated model of the group of thermostat-based devices, and a stochastic performance analysis.

### REFERENCES