Aggregate capacity of TCLs with cycling constraints

Austin R. Coffman
University of Florida
Gainesville, FL
bubbaroney@ufl.edu

Neil Cammardella
University of Florida
Gainesville, FL
cammadella@ufl.edu

Prabir Barooah
University of Florida
Gainesville, FL
pbarooah@ufl.edu

Sean Meyn
University of Florida
Gainesville, FL
meyn@ece.ufl.edu

ABSTRACT

A flexible load can vary its power consumption to perform grid support services. This flexibility is naturally limited by the Quality of Service (QoS) requirements at the load. A widely examined class of flexible loads is Thermostatically Controlled Loads (TCLs), which include air conditioners, water heaters, and refrigerators. A TCL is designed to maintain a temperature within a preset band, and the actuation to achieve this is on/off. Temperature, cycling rate, and the energy bill are three main QoS metrics: exceeding the temperature limits, frequent cycling between on and off, and a high energy bill must be avoided.

How the temperature constraint affects the capacity of an ensemble of TCLs to provide grid support is a well studied problem. However, how the cycling constraint effects the capacity is often neglected. In this work we present a characterization of the capacity of a collection of TCLs that takes into account not only temperature, but also cycling and energy constraints. Our characterization of capacity is consistent with its most practical utility: a grid authority can use this characterization to plan a reference signal that the TCLs can track without violating any of their QoS constraints. Additionally, the proposed characterization is independent of the algorithm used to coordinate the TCLs (to provide grid support) and leads to a convex and feasible optimization problem for the grid authority’s reference planning.

KEYWORDS

Virtual Energy Storage, Demand Dispatch, Capacity, Flexible Loads

1 INTRODUCTION

Falling prices and environmental stewardship have led to rapid growth of renewable energy resources. However, their inherent volatility creates new challenges for grid operators who must continuously balance the supply and demand of electric power. Currently, balance is maintained mostly through supply-side actions, i.e. generators are ramped up and down to meet demand. However, the ramping ability of generators is limited so that ancillary services are required to maintain power balance. It is possible to rely on fossil fuel or battery based ancillary services to maintain balance, but with downsides of negative environmental impact and cost, respectively. Motivated by these drawbacks a new environmentally friendly and cost effective resource has been investigated: Flexible Loads.

Flexibility refers to the ability of a load to deviate from a baseline level of power consumption without violating the Quality of Service (QoS) of the load. The grid operator or balancing authority (BA) would enable flexible loads by requesting the flexible loads to consume more or less power over baseline power consumption. Baseline power consumption refers to the power consumption that would have occurred without the BA interfering. From the perspective of the BA, this increase and decrease of consumption is identical to the charging and discharging of a battery. Due to this similarity, flexible loads assisting the grid are often termed Virtual Batteries (VB) [3]. From a cost perspective relying on VBs is financially more viable than actual batteries [7].

The term Virtual Energy Storage (VES) [3] is commonly used to denote flexible loads providing grid support services, while Demand Dispatch (DD) [4, 11] refers to the act of a grid authority dispatching the flexible loads to meet its needs. Unlike traditional demand side grid support services, which may only be used during extreme events, the vision for VES/DD is continuous operation to maintain power balance in the grid. Some examples of flexible loads that are suitable for VES/DD are Thermostatically Controlled Loads (TCLs) [5, 6, 11, 14, 19, 23, 26], HVAC systems in commercial buildings [17], and electric pumps for irrigation [1] and pool cleaning [16].

For flexible loads providing VES, the difference between the power consumption the BA requires and the baseline power consumption is the reference signal for the flexible load(s). In this work, we focus only on collections of TCLs so that the reference signal represents the requested amount of power deviation for an entire collection of TCLs.
The design of a coordination algorithm to control an ensemble of TCLs to track a given reference is a well studied problem in the academic literature [11, 13, 20, 23, 26]. However, while the BA is presumably aware of its needs, the BA must still determine how to allocate a portion of its needs to a collection of TCLs. In order for a BA to do this, the “capacity” of the collection of TCLs must be known. Put another way, without knowledge of the capacity the BA cannot effectively integrate TCLs into a DD program.

There is no agreed upon formal definition of capacity in the literature, other then conceptually the capacity represents limitations in aggregate behavior due to constraints at the individual TCL. The constraints are the QoS requirements of the individual TCL, which include: (i) the user’s thermal comfort (temperature) (ii) as TCLs are on/off loads, cycling and (iii) the user’s energy bill.

A requirement for the characterization of capacity is that it must be complete in the sense that it must account for all of the QoS constraints of the individual TCLs. If a BA constructs a reference signal with an incomplete notion of capacity, either: (i) the ensemble of TCLs will not be able to track the reference, or (ii) tracking the constructed reference will require the TCLs to violate their individual QoS requirements. In both scenarios the outlook of TCLs providing VES in the long term is grim; the BA views TCLs as an unreliable resource or the home owners (TCLs) view the BA as an authoritative monarch with unrealistic expectations.

Another requirement for the capacity characterization is that it should easily allow for a BA to perform aggregate level reference planning and resource allocation. Conceptually, one way to achieve this is to abstract the constraints at the individual TCLs to constraints on aggregate level quantities, such as the aggregate power deviation. Consistent with past literature, we term this abstraction as the characterization of the capacity and the resulting constraints as the aggregate capacity constraints.

Characterization of the capacity is challenging and a significant amount of research has been aimed at this [12, 15, 18, 25, 27, 28]. Most commonly characterizations specify constraints on the aggregate power and thermal energy deviation, which are aggregate versions of each individual TCL’s on/off and temperature state, respectively. However, much of the past literature is focused on characterizations that only account for the individual TCL’s temperature constraint [18, 27]. These characterizations are incomplete, and thus do not accurately describe the capacity.

Some works characterize the capacity to account for the individual TCLs cycling constraint [12, 15, 25, 28]. In our past work [15], to account for the cycling constraint of the individual, an algorithm dependent notion of capacity is developed. One disadvantage, in addition to being algorithm dependent, is that the proposed method in [15] requires the solution of a non-convex optimization problem, which is not ideal for real time planning. In [28] aggregate constraints are developed to account for the cycling constraint of the individual TCL, but reference planning is not done. Furthermore, the constraint on the aggregate power deviation is developed within the framework of a priority stack controller, making the capacity dependent on the coordination algorithm. Similar to [28] the cycling constraint in [12] is also coordination algorithm dependent. In [25], a centralized approach to handling aggregate capacity for TCLs with individual cycling constraints is taken. The authors furthermore provide a ramp rate constraint on the power deviation reference signal, to account for the individual TCL’s cycling constraint. However, using the constraints for reference planning requires a simulation of TCLs, making the capacity dependent on the coordination algorithm. In addition to TCLs, there is work on characterizing the capacity of deferrable loads [21].

In summary, the surveyed works that characterize the aggregate capacity of collections of TCLs suffer from a number of limitations: they account for only a subset of the individual’s QoS, and either (i) depend on the coordination algorithm or not convenient for BA-level reference planning. Thus, these characterizations do not meet the two aforementioned requirements.

In this work we characterize the capacity of a collection of TCLs as constraints on the aggregate power and thermal energy deviation. Our work is novel in three regards. First, the constraints are constructed to account for all three individual QoS: temperature, cycling, and monthly energy use. Second, the characterization is independent of the coordination algorithm used to control the ensembles of TCLs. Third, the capacity characterization can be used by a BA to compute the reference for an ensemble of TCLs by solving an optimization problem that is always feasible and convex. Together, these facets ensure that the reference signal so planned can be tracked with any well designed coordination algorithm that respects all three QoS constraints of each TCL.

The effectiveness of our capacity characterization is investigated in simulation experiments. An ensemble of TCLs are then coordinated with a priority stack controller, a modified version of the one developed in [18] (so as to enforce device cycling constraints), to track the computed reference. We offer a comparison of reference planning and tracking when the reference signal is planned using the aggregate capacity constraints of [18], which do not include information on individual TCLs cycling and energy QoS. The results of the comparison confirm the need to include all relevant individual TCL QoS constraints in reference planning.

The paper proceeds as follows: Section 2 describes the individual TCLs QoS and models, Section 3 describes quantities for N TCLs, Section 4 describes the derived aggregate capacity constraints, Section 5 and Section 5.2 describes the proposed and alternative reference planning methods, respectively, and Section 6 reports the results of the numerical experiments.

2 THE INDIVIDUAL TCL

2.1 QoS constraints

An on/off TCL is any device that turns on or off to maintain a temperature within a preset temperature deadband. Here, we denote the state space of a TCL as $X$, and elements of $X$ as the couple $x = \{m \in \{0, 1\}, \theta \in \mathbb{R}\}$, that consists of the off (0) and on (1) status (mode) and temperature of the TCL. We denote the electrical power consumption as $P$, which is the power consumed by the TCL when it is on. Furthermore, time is discrete and denoted by the index $k$ with total time horizon $N_k$ and the TCL index is $j$. With the state variables declared, operating constraints (QoS) for the $j^{th}$ TCL are
Aggregate capacity of TCLs with cycling constraints

We denote the set \( Q_{\delta}^k \triangleq \{ \theta_{\text{set}}, \delta, \tau_{\text{ci}}, \bar{E}, N_k \} \) as the "QoS set," which is the set that contains the user defined parameters that appear in (1)-(3). The variables \( \theta_{\text{set}} \) (set point) and \( \delta \) specify the temperature deadband: \( \theta_{\text{min}} \leq \theta_{\text{set}} \leq \theta_{\text{max}} \). The full width temperature deadband is denoted as \( \Delta \triangleq \theta_{\text{max}} - \theta_{\text{min}} \). The variable \( \tau_{\text{ci}} \) is the parameter for the cycling constraint (2). The variable \( N_k \) (\( N_k \leq N_i \)) is the length of the time interval during which the energy deviation is evaluated, and \( \bar{E} \) (\( \bar{E} \geq 0 \)) represents the permitted energy deviation during that interval. The quantity \( \bar{P}_k^j \) is the baseline power consumption at time slot \( k \) and \( T_s \) is the sample time.

The first constraint is that the temperature remain within the temperature deadband. The second constraint is that the device can only switch once within a specified period \( \tau_{\text{ci}} \). The third constraint is that the time averaged energy deviation during a time interval of length \( N_k \) not deviate from baseline energy consumption by a pre-specified amount. Practically, \( N_k \) can represent the length of an electricity billing period so that constraint (3) is a constraint to keep the energy bill close to what the user pays during baseline consumption.

**Definition 1.** A TCL switches on (respectively, off) at time \( k \) if at time \( k \) it is on (respectively, off) and at time \( k-1 \) it is off (respectively, on).

We represent switch on and off as the variables, \( S_{k-1}^{\text{on},j} \) and \( S_{k-1}^{\text{off},j} \), defined as,

\[
S_{k-1}^{\text{on},j} \triangleq \begin{cases} 1, & \text{if } \left| m_k - m_{k-1} \right| = 1. \\ 0, & \text{otherwise.} \end{cases}
\]

\[
S_{k-1}^{\text{off},j} \triangleq \begin{cases} 1, & \text{if } \left| m_{k-1} - m_k \right| = 1. \\ 0, & \text{otherwise.} \end{cases}
\]

An on off switch can occur because of two events: (i) the TCL switches to maintain the temperature QoS (1) or (ii) the TCL switches for the purpose of providing VES:

\[
S_{k-1}^{\text{on},j} = F_{k-1}^{\text{on},j} + D_{k-1}^{\text{on},j},
\]

\[
S_{k-1}^{\text{off},j} = F_{k-1}^{\text{off},j} + D_{k-1}^{\text{off},j}.
\]

The quantity \( F_{k-1}^{\text{on},j} \) (respectively, \( F_{k-1}^{\text{off},j} \)) represents the on switch to provide VES (respectively, off switch). The quantity \( D_{k-1}^{\text{on},j} \) (respectively, \( D_{k-1}^{\text{off},j} \)) represents a switch to maintain the temperature QoS (1).

**Definition 2.** A TCL is stuck off (respectively, on) if it is off (respectively, on) and has changed mode once in the past \( \tau_{\text{ci}} \) times.

We represent stuck on and off as the variables, \( y_{k}^{\text{on},j} \) and \( y_{k}^{\text{off},j} \), defined as,

\[
y_{k}^{\text{on},j} \triangleq \begin{cases} 1, & \text{if } \sum_{i=0}^{\tau_{\text{ci}}-1} \left| m_{k-i} - m_{k-i-1} \right| = 1, m_k = 1. \\ 0, & \text{otherwise.} \end{cases}
\]

\[
y_{k}^{\text{off},j} \triangleq \begin{cases} 1, & \text{if } \sum_{i=0}^{\tau_{\text{ci}}-1} \left| m_{k-i} - m_{k-i-1} \right| = 1, m_k = 0. \\ 0, & \text{otherwise.} \end{cases}
\]

**2.2 Modeling and Control of the individual TCL**

As in much of prior work [13, 23] temporal evolution of the temperature \( \theta_k^j \) is modeled in discrete time as a linear difference equation

\[
\theta_{k+1}^j = a \theta_k^j + (1 - a) \left( \theta_k^a - \bar{R}_{th} m_k Q_{ac} \right),
\]

\[
\bar{a} = \exp \left( \frac{-T_s}{R_{th} C_{th}} \right),
\]

where \( R_{th} \) and \( C_{th} \) represent the thermal resistance to ambient temperature \( \theta_k^a \) and thermal capacitance of \( \theta_k^j \), respectively. The thermal power consumption \( Q_{ac} \) is related to the electrical power consumption by \( Q_{ac} = \eta P \), where \( \eta \) is the Coefficient of Performance (COP). The thermal energy deviation quantity [18] of the \( j \)th TCL is defined as,

\[
z_k^j \triangleq C_{th} (\theta_k^j - \theta_{\text{set}}),
\]

which is an affine transformation of the temperature of the TCL, \( \theta_k^j \). The relationship with electrical energy storage is explored in [8].

The dynamics for thermal energy are obtained by substituting the definitions for \( z_k^j \) and \( z_{k+1}^j \) into (10),

\[
z_{k+1}^j = \bar{a} z_k^j - b \left( m_k^j \bar{P}_k^j - \frac{\theta_k^a - \theta_{\text{set}}}{\eta R_{th}} \right),
\]

\[
b = (1 - \bar{a}) C_{th} R_{th}.
\]

We identify the RHS term in parenthesis in (13) as the power deviation for the \( j \)th TCL, so that the baseline power consumption for the \( j \)th TCL is,

\[
\bar{P}_k^j = \frac{\theta_k^a - \theta_{\text{set}}}{\eta R_{th}},
\]

This form of the baseline power consumption is a consequence of the equation (10) used to model the TCL. This quantity can be used, e.g. to evaluate (3).

**2.2.1 Coordination of TCLs to provide VES**

Typically, a local thermostatic controller is responsible for enforcing the individual TCLs QoS (1)-(3). However, when TCLs provide VES the role of the thermostat is subsumed by the coordination algorithm that is subsequently required. Here the algorithm simultaneously enforces each TCLs QoS and coordinates the ensemble is irrelevant and independent to our work. In what follows, aggregate quantities and constraints will be developed from the individual quantities and constraints in this section. A BA can then use the constraints on the aggregate quantities to plan aggregate power deviation trajectories.
for a coordination algorithm. The key being that the planned trajectories allow for any well designed coordination algorithm to enforce QoS constraints (1)-(3) while maintaining tracking performance.

3 AGGREGATE QUANTITIES

Section 2 was devoted to the individual TCL; we now define variables for a collection of N TCLs. Two quantities are of interest at the aggregate level: (i) total quantities in units of power (Watts) and (ii) fractional quantities normalized by $N$, the number of TCLs. Furthermore, a homogeneous collection is defined as an ensemble of TCLs for which the parameters that appear in (10) ($C_{th}, R_{th}, \eta, P$) and the QoS set $Q_{th}$ are uniform over the population. A homogeneous collection of TCLs is considered, in the following.

The total power consumption of the collection at time $k$ is

$$Y_k \triangleq N^{\text{on}}_k P = P \sum_{j=1}^{N} m^j_k$$

where $N^{\text{on}}_k$ is the number of TCLs on at time $k$; the number of TCLs off at time $k$ is $N^{\text{off}}_k = N - N^{\text{on}}_k$. The aggregate thermal energy deviation is defined as,

$$\gamma_k \triangleq \frac{1}{N} \sum_{j=1}^{N} \gamma^j_k$$

The maximum aggregate power consumption is,

$$P_{\text{agg}} \triangleq NP,$$

which is the power consumption that occurs when all of the TCLs are on. Another important aggregate quantity to be defined is the baseline power consumption. The baseline for the ensemble is,

$$P_{\text{off}} \triangleq \sum_{j=1}^{N} \rho^j_k = N \left( \frac{\theta^j_k - \theta_{\text{set}}}{\eta R_{th}} \right).$$

The fractional quantities are denoted,

$$\gamma^{\text{on}}_k \triangleq \frac{\sum_{j=1}^{N} \gamma^{\text{on},j}_k}{N}, \quad \gamma^{\text{off},j}_k \triangleq \frac{\sum_{j=1}^{N} \gamma^{\text{off},j}_k}{N},$$

$$\rho^{\text{on}}_k \triangleq \frac{N^{\text{on}}_k}{N}, \quad \rho^{\text{off}}_k \triangleq \frac{N^{\text{off}}_k}{N},$$

$$d^{\text{on}}_k \triangleq \frac{\sum_{j=1}^{N} D^{\text{on},j}_k}{N}, \quad d^{\text{off}}_k \triangleq \frac{\sum_{j=1}^{N} D^{\text{off},j}_k}{N},$$

$$y^{\text{on}}_k \triangleq \frac{\sum_{j=1}^{N} Y^{\text{on},j}_k}{N}, \quad y^{\text{off}}_k \triangleq \frac{\sum_{j=1}^{N} Y^{\text{off},j}_k}{N}.$$ (23)

The aggregate power deviation, over baseline power consumption is denoted

$$y_k \triangleq P_{\text{agg}} \gamma^{\text{on}}_k - P_{\text{off}}.$$ (24)

The power deviation reference signal for a collection of TCLs is,

$$r_k \triangleq \text{Desired value of } y_k \text{ at time } k.$$ (25)

**Comment 1.** For a homogeneous collection of TCLs, the fraction of loads on and the total power consumption are proportional. Thus in the developments to follow, “fraction of loads on” and “total power consumption” can be freely interchanged, modulo a scaling factor.

4 AGGREGATE CAPACITY CONSTRAINTS FOR REFERENCE PLANNING

Aggregate capacity constraints refers to constraints on aggregate quantities due to constraints at the individual TCL, e.g. (1)-(3). We formulate constraints on aggregate quantities of the two individual TCL states (i) power deviation (24) and (ii) thermal energy deviation (17). Our constraints on aggregate power and thermal energy deviation account for the temperature, cycling, and energy constraint at the individual, specified by (1)-(3), respectively. That is, these aggregate constraints ensure that if a power and thermal energy deviation trajectory were to satisfy them, then a collection of TCLs could track the power deviation signal while enforcing (1)-(3). Conversely, if the aggregate constraints are violated, then there would exist at least a single TCL that violates its individual QoS constraints.

4.1 Fraction Stuck

The fraction of TCL’s stuck on, or off, can be represented as an inventory model [2, 24] with deterministic demand,

$$Y_k = Y_{k-1} + B \gamma_{k-1} - \gamma_{k-1} \tau_{\text{BA}}.$$ (26)

In words, the fraction that are stuck on, $Y^{\text{on}}_{k-1}$ (respectively, off), is increased by the fraction that switch on $\gamma^{\text{on}}_{k-1}$ (respectively, off) from $k-1$ to $k$ and decreased by the fraction that had switched on (respectively, off) $k-1 - \tau_{\text{BA}}$ sample times in the past. A derivation for (26) is given in the appendix. We define an input for stuck on (respectively, off) as the following column vector,

$$i^{\text{on}}_{k-1} = [i^{\text{on}}_{k-1}, \ldots, i^{\text{on}}_{k-1 - \tau_{\text{BA}}}]^T,$$

the $T$ superscript denotes matrix transpose. Eq. (26) can now be represented as follows, which is a linear state space model,

$$y_k^{\text{on}} = y_{k-1}^{\text{on}} + B(\tau_{\text{BA}})u_{k-1}^{\text{on}}, \quad y_0^{\text{on}} = 0,$$ (28)

$$y_k^{\text{off}} = y_{k-1}^{\text{off}} + B(\tau_{\text{BA}})u_{k-1}^{\text{off}}, \quad y_0^{\text{off}} = 0.$$ (29)

For both systems, the matrix $B(\tau_{\text{BA}})$ is

$$B(\tau_{\text{BA}}) \triangleq \begin{bmatrix} 1 - 0^{\text{BA}} - 1 \end{bmatrix},$$ (30)

where $\theta_s$ is a row vector of zeros of length $r$. The quantity $\tau_{\text{BA}}$ is elected as $\tau_{\text{BA}} > \tau_{\text{ref}}$, and is the cycling QoS parameter the BA uses for reference planning.

4.1.1 $\tau_{\text{BA}} > \tau_{\text{ref}}$. TCLs may have lockout times as little as 5 minutes [9], however this does not mean it is desirable for a TCL to switch every 5 minutes. So that using $\tau_{\text{BA}} > \tau_{\text{ref}}$ will allow the BA to plan a reference signal that would require TCLs to switch less over a given time horizon.

4.2 Power Deviation Limits

We start by considering how much the fraction of on devices could be changed in a given sample time, relative to the current fraction of on devices. To obtain an upper bound on the change $n^{\text{on}}_k - n^{\text{on}}_{k-1}$, assume that $n^{\text{on}}_k \geq n^{\text{on}}_{k-1}$. The quantity $n^{\text{on}}_k - n^{\text{on}}_{k-1}$ represents the current fraction of TCLs that are off and can switch on, so that the upper bound on $n^{\text{on}}_k - n^{\text{on}}_{k-1}$ should include at least:

$$i^{\text{off}}_{k-1} = i^{\text{off}}_{k-1}.$$
However, this is not complete as some TCLs may be forced to switch due to the temperature constraint (1). The upper bound should then be increased by $d^\text{on}_{k-1}$ and decreased by $d^\text{off}_{k-1}$. Letting $\Delta k_{-1} := d^\text{on}_{k-1} - d^\text{off}_{k-1}$ an upper bound is

$$n^\text{on}_k \leq 1 - \gamma^\text{off}_k + \Delta k_{-1},$$

(31)

where $\gamma^\text{on}_k$ is eliminated through the relation $\gamma^\text{off}_k = 1 - n^\text{on}_k$. The steps necessary to obtain the lower bound are symmetric, and the result is

$$\gamma^\text{on}_{k-1} + \Delta k_{-1} \leq n^\text{on}_k \leq 1 - \gamma^\text{off}_k + \Delta k_{-1}.$$

(32)

Practically, the quantity $\Delta k_{-1}$ is small in magnitude. Unless the ambient conditions are extreme, due to randomness of initial conditions, the fraction of TCLs that switch at on at $k$ due to hitting the upper bound of the temperature deadband, $d^\text{on}_{k-1}$, should be close to the fraction of TCLs that switch at off at $k$ due to hitting the lower bound of the temperature deadband, $d^\text{off}_{k-1}$, which leads to $\Delta k_{-1} = 0$ since $\Delta k_{-1} = d^\text{on}_{k-1} - d^\text{off}_{k-1}$. We then remove this quantity from the bound (32) to obtain,

$$\gamma^\text{on}_{k-1} \leq n^\text{on}_k \leq 1 - \gamma^\text{off}_k.$$

(33)

Removal of $\Delta k_{-1}$ is an approximation, and thus measures must be taken so that reference signals designed with (33) do not cause significant tracking errors. We believe the following will help mitigate tracking errors: $\tau_{\text{BA}} > \tau_{\text{rcl}}$ where $\tau_{\text{rcl}}$ and $\tau_{\text{BA}}$ are the cycling QoS parameters for individual TCLs and the BA, respectively (described near (28)-(29)).

When a BA designs a reference signal with $\tau_{\text{BA}} > \tau_{\text{rcl}}$, it is underestimating the capacity of the collection. That is, the BA is assuming that TCLs can switch less than they actually can. The hope is that when $\Delta k_{-1}$ contributes in a non-conservative manner to (32), the extra capacity available due to $\tau_{\text{BA}} > \tau_{\text{rcl}}$ will enable the collection to counteract the effect of $\Delta k_{-1}$ and continue to track the reference signal designed with (33).

Comment 2. Past work has either identified the quantity $\Delta k_{-1}$ as a challenge [25], ignored it [28], or developed weaker bounds independent of $\Delta k_{-1}$ [15].

4.3 Thermal Energy Deviation Limits

The dynamics for the aggregate thermal energy are obtained by summing (13) over the index $j$, which results in

$$z_{k+1} = az_k - br_k,$$

(34)

where $r_k$ is the power deviation reference trajectory of the ensemble of TCLs, $a$ is as defined in (11), and $b$ is as defined in (14).

In the past literature [18], the bounds for aggregate thermal energy (for a homogeneous collection) is obtained as follows,

$$|\theta^j_k - \theta^\text{set}_k| \leq \delta \Rightarrow |z^j_k| \leq C = \frac{\gamma^\text{th} A}{2\eta}, \quad \forall k,$$

(35)

so that, with (35) and the triangle inequality,

$$|z_k| = \sum_{j=1}^{N} |z^j_k| \leq NC, \quad \forall k.$$

(36)

However, when individual TCLs have cycling constraints the bound (35) and consequently (36) are incorrect when the cycling constraint (2) is “active,” i.e. the TCL is stuck on or off. In what follows, we adopt the bound (35) as the thermal energy bound for a TCL that is not stuck on or off.

The need for a new thermal energy bound is explained as follows: if a cooling TCL is stuck on at $k$ then the temperature of that TCL cannot increase at $k + 1$. Contrary, if a cooling TCL is stuck off at $k$, the temperature cannot decrease at $k + 1$. This inability to increase or decrease temperature is what causes (2) to induce a bound on the thermal energy deviation of the individual. Thermal energy deviation is a scaled version of temperature deviation (12).

If TCL $j$ is stuck on at time $k$ (which implies thermal energy can only decrease) then the right upper bound is $z_k^j$ and the lower bound is unaffected ($-C$). Contrary if TCL $j$ is stuck off (which implies thermal energy can only increase) then the right lower bound is $z_k^j$ and the upper bound is unaffected ($C$). Due to this asymmetry, we consider a separate upper and lower bound on the aggregate thermal energy deviation. The bounds are constructed by summing the bounds for the individual TCLs thermal energy deviation and separating the TCLs that are stuck on (respectively, off) in the upper bound (respectively, lower bound). The upper and lower bounds are denoted respectively,

$$C_k^+ = N(1 - \gamma^\text{on}_k C) + \sum_{j \text{ stuck on}} z^j_k,$$

(37)

$$C_k^- = -N(1 - \gamma^\text{off}_k C) + \sum_{j \text{ stuck off}} z^j_k,$$

(38)

The summation in $C_k^+(C_k^-)$ is over loads stuck on (stuck off), where the bound $C_k^- \leq z_k \leq C_k^+$ follows.

In order to compute the bounds (37) and (38), the individual thermal energy deviation, $z^j_k$, would have to be known for all of the TCLs that are stuck on or off. Since it is intractable to know the thermal energy deviation of the $j^{th}$ TCL, we make a worst case approximation. This approximation assumes that the portion of TCL’s that are stuck on or off are stuck at the extreme limit of the deadband, i.e. a lower or upper bound for the two summation terms,

$$-NC \gamma^\text{on}_k \leq \sum_{j \text{ stuck on}} z^j_k,$$

(39)

$$NC \gamma^\text{off}_k \geq \sum_{j \text{ stuck off}} z^j_k,$$

(40)

So that new thermal energy deviation bounds are derived,

$$\tilde{C}_k^+ \triangleq NC(1 - 2\gamma^\text{on}_k) \leq C_k^+, \quad \text{and} \quad \tilde{C}_k^- \triangleq -NC(1 - 2\gamma^\text{off}_k) \geq C_k^-,$$

(41)

(42)

and the thermal energy deviation satisfies,

$$\tilde{C}_k \leq z_k \leq \tilde{C}_k.$$

(43)

This implies the bounds (37)-(38) by construction.

4.4 Relation to “fraction on”

The fraction stuck on $\tilde{\gamma}^\text{on}_k$ (respectively, off $\tilde{\gamma}^\text{off}_k$) is related to the fraction of on (respectively, off) switches through the dynamics (28)
Another “inventory equation” couples switching and power model dynamics:
\[ n_{k}^{\text{on}} = n_{k-1}^{\text{on}} + n_{k-1}^{\text{on}} - s_{k-1}^{\text{off}}. \]

In words, the fraction of on devices at time \( k \) is the fraction on at \( k-1 \), plus the fraction that switch to on (\( s_{k-1}^{\text{on}} \)) and minus the fraction that switch to off (\( s_{k-1}^{\text{off}} \)) from time step \( k-1 \) to time step \( k \). A derivation for (44) is in the appendix. For notational consistency, we re-write (44) as,
\[ n_{k}^{\text{on}} = n_{k-1}^{\text{on}} + u_{k-1}^{\text{on}}[1] - u_{k-1}^{\text{off}}[1], \]

where \( u_{k-1}^{\text{on}}[1] \) represents the first element of the vector \( u_{k-1}^{\text{off}} \), i.e. \( u_{k-1}^{\text{on}}[1] = s_{k-1}^{\text{off}} \).

**Comment 3.** The relationship between the fraction of devices on and the fraction of on and off switches (44) is a quantity independent from the coordination algorithm. Meaning, regardless of how the population of TCLs is controlled, the fraction on can be thought of as a dynamic discrete time system with the fractional switching differential as input.

### 4.5 VES Constraint

The BA requires one constraint to ensure that the collection of TCLs do not act as generators, namely:
\[ \sum_{k=0}^{N_{k}-1} r_{k} = 0. \]

We now show that this constraint is a necessary condition for the individual TCLs energy constraint (3). We assume that \( N_{k} = N_{b} \), which loses no generality as \( N_{k} \) is arbitrary and would already be a function of \( N_{b} \). Summing (3) over the \( j \) index and expanding the absolute value,
\[ -N \sum_{j=1}^{N_{k}} E_{j} \leq \frac{T_{k}}{N_{k}} \sum_{j=1}^{N_{k}} \sum_{k=0}^{N_{k}-1} (m_{j}^{k} p - P_{j}^{k}) \leq \sum_{j=1}^{N_{k}} E_{j}. \]

Converting back to absolute value, the aggregated version of (3) is
\[ \frac{T_{k}}{N_{k}} \sum_{k=0}^{N_{k}-1} r_{k} \leq \sum_{j=1}^{N_{k}} E_{j}, \]

which due to (46) will be true for all values of \( E_{j} \), as the RHS term in (50) is defined to be greater than or equal to zero. If (50) is not satisfied, then it can be shown through the law of the contrapositive that there would exist at least a single TCL that does not satisfy (3). In the scenario that the individual TCLs do not have symmetric energy constraints, then the aggregate version of (3) would resemble (49); The constraint (46) still enforces this.

## 5 Reference Planning

The reference planning problem utilizes the constraints developed in Section 4 so to plan a power deviation trajectory (reference signal) for an ensemble of TCLs. The goal of the reference planning problem is to project the BA’s total desired demand deviation, \( r_{BA}^{k} \), onto the set defined through the aggregate capacity constraints. The signal \( r_{BA}^{k} \) can be a regulation signal, or it can be obtained by filtering the net demand [3], though a discussion is outside the scope of this work.

### 5.1 Proposed method

We collect the “battery model” (34) with the aggregate power deviation constraint (33) and thermal energy deviation constraint (43) to form an aggregate constraint set. The aggregate power (33) and thermal energy (43) deviation constraints are coped to the variables of the battery model (34) through (44). The full length decision vector for the constraint set is,
\[ \psi \triangleq \{ (z_{k})_{1}^{N_{k}}, (r_{k})_{0}^{N_{k}-1}, (u_{k}^{\text{on}})_{1}^{N_{k}-1}, (u_{k}^{\text{off}})_{0}^{N_{k}-1}, \ldots \}
\]
\[ \{ (r_{k})_{0}^{N_{k}-1}, (y_{k}^{\text{on}})_{1}^{N_{k}-1}, (y_{k}^{\text{off}})_{0}^{N_{k}-1} \}. \]

The constraint set at time \( k \) for the ensemble of TCLs, based on the aggregate capacity constraints developed in Section 4, is:
\[ \Omega_{k} \triangleq \left\{ \psi_{k} = (z_{k}, r_{k}, u_{k}^{\text{on}}, u_{k}^{\text{off}}, y_{k}^{\text{on}}, y_{k}^{\text{off}}) : \right. \]
\[ \left. \begin{array}{l}
    z_{k} = a z_{k-1} - b r_{k-1}, \quad z_{0} = z, \\
    \frac{\gamma_{k} C_{k}}{\gamma_{k} C_{k} + \tilde{C}_{k}} z_{k} \leq \tilde{C}_{k}, \\
    \frac{1}{P_{\text{agg}}} \left( r_{k} + \bar{p}_{k} \right), \\
    y_{k}^{\text{on}} \leq y_{k}^{\text{off}}, \quad y_{k-1}^{\text{off}} = 0, \quad y_{k-1}^{\text{on}} = 0; \\
    y_{k}^{\text{on}} = y_{k-1}^{\text{on}} + B(r_{BA}^{k}) u_{k}^{\text{on}} - y_{k}^{\text{off}}; \\
    y_{k}^{\text{off}} = y_{k-1}^{\text{off}} + B(r_{BA}^{k}) u_{k}^{\text{off}}, \\
    \psi_{k} \in \Xi_{k}, \quad \Xi_{k} := \sum_{k=0}^{N_{k}-1} r_{k} = 0, \quad N_{k} \right\}. \]

The values \( n_{\psi} \) and \( z_{\psi} \) are the initial conditions. The reference planning is achieved by projecting a reference signal known to the BA, \( r_{BA}^{k} \), onto the aggregate constraint set. We define the full length projection vector as,
\[ \psi^{BA} \triangleq \{ (0)_{1}^{N_{k}}, (1)_{1}^{N_{k}} r_{BA}^{k}, (0)_{1}^{N_{k}}, (1)_{1}^{N_{k}} y_{BA}^{k}, (0)_{1}^{N_{k}}, (1)_{1}^{N_{k}} \}. \]

so that the optimization problem to be solved is,
\[ \min_{\psi} \left\{ \sum_{k=0}^{N_{k}-1} (\psi^{BA} - \psi_{k})^{T} \Xi (\psi^{BA} - \psi_{k}) \right\} \]
\[ \text{s.t.} \quad \forall k \in \{1, \ldots, N_{k}\}, \quad \psi_{k} \in \Omega_{k}, \quad \sum_{k=0}^{N_{k}-1} r_{k} = 0. \]
LEMMA 1. The objective function (61) is strictly convex, the inequality constraints are feasible, and the set $\bigcap_{k=1}^{N_f} \Omega_k$ is non-empty and convex.

The proof of Lemma 1 is given in the Appendix. A consequence of Lemma 1 is that the solution to this problem exists and is unique. In other words, for any $\psi^{BA}$ there will always exist a unique reference signal that a collection of TCLs are ideally suited to track.

COMMENT 4. From the proof of Lemma 1 the vector $\psi = 0$ is in the total constraint set. So setting the “projection” elements, excluding $r_k^{BA}$, of $\psi^{BA}$ to zero is equivalent to desiring the other decision variables to “stay” within the set. Practically, the solution obtained will best track $r_k^{BA}$ while also minimizing the fraction of on/off switches, the fraction stuck on/off, and the aggregate thermal energy. The relative magnitude of the diagonal elements in $\Xi$ specify the level of preference for each of these goals; the decision variable that corresponds to the largest value in $\Xi$ will take precedence in minimizing its distance to the corresponding element in $\psi^{BA}$.

5.2 Alternative Method

To compare with past literature we define a constraint set based on the constraints developed in [18] and the battery model (34) for projection of $r_k^{BA}$. The disadvantage with this constraint set is that the aggregate power and thermal energy deviation bounds developed in [18] do not account for the individual cycling (2) or energy (3) constraint. This alternative reference planning problem is posed as,

$$\min_{(r_k)_{k=1}^{N_f-1},(z_k)_{k=1}^{N_f}} \Xi[2,2] \sum_{k=0}^{N_f-1} (z_k^{BA} - r_k)^2 + \Xi[1, 1] \sum_{k=1}^{N_f} z_k^2$$

s.t. $\forall k \in \{0, ..., N_f - 1\}$

$$z_{k+1} = a z_k - b r_k, \quad z_0 = z,$$

$$(65)$$

$$|z_k| \leq N_C, \quad -P_k \leq r_k \leq P_{agg} - \tilde{P}_k.$$  \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} \hspace{1cm} (66)

The value $\Xi[i, j]$ represents the $(i, j)^{th}$ entry of the matrix $\Xi$ in (61), so the objective function weighting for both methods is consistent. If compared to the bounds developed in Section 4, the bounds for $z_k$ and $r_k$ in (66) assume that no TCLs are stuck on or off. This is a particularly generous assumption, as TCLs (at least air conditioner TCLs) have local cycling constraints that render them stuck on or off [9].

6 NUMERICAL EXPERIMENTS

We conduct two numerical experiments: (i) comparison of our proposed reference planning method to the alternative method posed in Section 5.2 and (ii) a parametric study to evaluate the effectiveness of using $T_{BA} > T_{rcl}$ to merit the assumption $\Delta d_k = 0$. For clarity, the simulated TCLs are residential air conditioner units (ACs). Additionally, all scenarios involve the solution of a convex optimization problem, which is performed using CVX [16]. For all experiments the sampling time is $T_s = 2$ minutes.

In the method comparison scenario, the purpose is to illustrate that all individual TCL QoS must be accounted for in reference planning. The proposed and alternative methods are used to plan two reference signals. We then use a priority stack controller to coordinate an ensemble of TCLs to track the planned reference signals. Under priority stack control, we present three tracking scenarios: (t-i) tracking the reference from the proposed method (t-ii) tracking the reference from the alternative method and (t-iii) tracking the reference from the alternative method with TCLs that disregard their cycling QoS. The intent being to show that only in scenario (t-i) will the ensemble of TCLs be able to track the planned reference while each individual enforces its own QoS.

The priority stack controller mentioned in the above scenarios is a modified version of the one presented in [18], so to also enforce the individual TCLs cycling (2) QoS; it by default enforces the temperature (1) QoS. The priority stack controller does not enforce the energy QoS (3), as selecting appropriate bounds $(\bar{E}^r)$ is a function of the price of electricity, which is out of the scope of this paper.

COMMENT 5. Since our capacity characterization is coordination algorithm independent, the choice of the coordination algorithm is arbitrary. However, we use a priority stack controller in the above mentioned experiments due to its simplicity. Coordination algorithms that depend on multitudes of tuning parameters could obscure the results if these hyper parameters are not tuned correctly.

6.1 Method Comparison: Reference Planning and Tracking

For both reference planning methods the BA supplied reference, $r_k^{BA}$, is obtained from BPA, a Balancing Authority in the Pacific Northwest of the United States, and is shown in Figure 1. The parameters for the individual loads are elected based on the values provided in [22] and are shown in Table 1, along with other simulation parameters.

Figure 1 shows the planned reference signals for both methods. The reference signal planned with the proposed method is noticeably less aggressive than the reference signal planned with the alternative method. That is, when cycling constraints are not taken into account higher ramp rates are asked from the collection of TCLs. As we will see briefly, this leads to either poor reference tracking, violation of individual TCLs QoS, or both.

In Figure 2 (top), the reference tracking results are shown for our proposed method that includes cycling information in reference planning. The priority stack controller is able to coordinate the collection of AC units to track the planned reference signal with minimal tracking error (see Table 2). For verification, the individual cycling QoS results are shown in Figure 2 (bottom). Every AC unit maintains to the preset level, as no units cycle faster than $T_{rcl} = 10$ minutes.
Figure 1: BA signal ($r_{BA}^k$) and the reference trajectories ($r_k$) for a collection of 60,000 TCLs.

Table 2: Reference Tracking Errors

| Reference planning method | Tracking Error |
|---------------------------|----------------|
| Proposed method           | 0.06 %         |
| Alternative method        | 21 %           |

In Figure 3 (top), the reference tracking results are shown for the alternative method that does not include cycling information in reference planning. Since this reference is beyond the capacity of the TCLs, and the priority stack controller enforces cycling QoS, it is unable to coordinate the collection of AC units to track the planned reference. For comparison, the reference tracking error reported in Table 2 is two orders of magnitude higher than the error with our proposed method. This illustrates that TCLs cycling constraints should be incorporated in reference planning.

Another consequence of the reference from the alternative method neglecting the capacity is that this actually prevents the priority stack controller from enforcing the cycling QoS, Figure 3 (bottom). The reference signal is requiring TCLs to switch on or off too close to the deadband, so that when a TCL switches to enforce (1) it will have switched in a time less than $\tau_{tcl}$ from its previous switch.

Figure 4 shows the results of tracking the reference signal planned from the alternative method when the priority stack controller does not enforce the cycling QoS (2) at the TCLs. At the cost of roughly 20 % of the total switches occurring 2 minutes apart (the sampling time), the tracking is near perfect. From experience this result is consistent across sampling times; The constraints in the alternative method assume the ability of the TCLs to switch at the sampling time.

6.2 Parametric Study

In this study $r_{BA}^k$ is the same as the method comparison scenario and the parameters are as shown in Table 1, except for $\tau_{tcl}$ and $\tau_{BA}$ which are varied over a range. To proceed with the parametric study over $\tau_{tcl}$ and $\tau_{BA}$, we define the metrics, $s^\tau$ and $d^\tau$. The first is, $s^\tau$:

$$s^\tau = \frac{1}{N} \sum_{j=1}^{N} s^j, \quad \text{where} \quad s^j = \frac{1}{N} \sum_{k=1}^{N_j} \left| m^j_k - m^j_{k-1} \right|, \quad (67)$$
which counts the total number of TCL switches normalized by the number of TCLs. The desired value of $s^r$ is small, as a smaller number of total switches is preferred, but some amount of switching is required for providing VES and maintaining thermal comfort (1).

The second metric is,

$$d^r \triangleq \sum_{k=1}^{N_t} H_k,$$

where the variable $H_k$ is defined as,

$$H_k \triangleq \begin{cases} 1, & (1 - y_{k-1}^{off} + \Delta d_{k-1}) < \frac{r_k + \Delta P_{agg}}{P_{agg}} \quad \text{and} \quad (r_k - y_{k-1}) > 0. \\ 1, & (y_{k-1}^{on} - \Delta d_{k-1}) > \frac{r_k + \Delta P_{agg}}{P_{agg}} \quad \text{and} \quad (y_{k-1} - r_k) > 0. \\ 0, & \text{otherwise}. \end{cases}$$

$H_k$ is 1 if the current fraction of TCLs on, $(y_{k-1} + \Delta P_{agg})/P_{agg}$, is required to increase or decrease to a value, $(r_k + \Delta P_{agg})/P_{agg}$, that is beyond the derived limits in (32) and 0 otherwise. Thus this metric counts the total number of times that the aggregate power capacity is exceeded. The desired value of $d^r$ is zero.

The two metrics are computed as follows: (i) use our proposed method and a given $\tau_{BA}$ value to plan a reference signal, $(r_k)_{k=0}^{N_t-1}$ (ii) use the priority stack controller to coordinate an ensemble of simulated TCLs with cycling metric $\tau_{c}$ to track $r_k$ and (iii) after the simulation collect the relevant data for computation of $s^r$ and $d^r$.

The values of the two metrics for various values of $\tau_{c}$ and $\tau_{BA}$ are shown in Figure 5, and the results indicate that when $\tau_{BA}$ is increased both $s^r$ and $d^r$ are decreased. In other words, the use of $\tau_{BA} > \tau_{c}$ is working as expected.

From the results of the parametric study it is also clear that $\tau_{BA} > \tau_{c}$ is necessary if $\Delta d_{c}$ is to be removed from the power bound (32). For instance, consider a scenario for which TCLs implement $\tau_{c} = 15$ (30 mins.). If the reference is designed with $\tau_{BA} = 15$ there are points in time where the capacity is exceeded (as indicated in Figure 5, bottom). Although, if the reference were designed with $\tau_{BA} = 40$ the capacity is never exceeded for $\tau_{c} = 15$.

The choice $\tau_{BA} > \tau_{c}$ was also included to reduce the total number of switches for a TCL. The success of this is documented in the top of Figure 5, as when $\tau_{BA}$ is increased $s^r$ is decreased. As another example, if the desired opt-out time is $\tau_{c} = 5$, designing a reference with $\tau_{BA} = 15$, instead of $\tau_{BA} = 5$, decreases $s^r$ from 22 to 10. Roughly, this means that the average TCL engages in half the amount of mode state switches over the same given time horizon.

### 7 CONCLUSION

We present an aggregate capacity characterization for collections of TCLs with individual temperature, cycling, and energy QoS constraints. We then use this characterization to pose the BA’s reference planning problem as an optimization problem, in which the power deviation signal desired by the BA is projected onto the set of signals that are within the ensemble’s capacity to track. This differentiates our approach from past literature that is largely focused on just temperature constraints, or are unsuitable for reference planning. Our aggregate capacity characterization takes the form of constraints that can be considered as necessary conditions for maintaining every individual TCL’s QoS: if not satisfied there will
be at least one TCL whose QoS requirements will not be met. This is then verified through simulations: when a reference is planned with an incomplete characterization of capacity that considers only temperature constraints, then tracking of this reference causes TCLs to cycle frequently; a behavior that would prevent homeowners from participating in grid support programs. This unwanted behavior is absent when the proposed capacity characterization is used for reference planning.

In the future we need to extend our work to handle time varying ambient conditions as well as heterogeneity among loads. Additionally, the determination of a relationship between $\tau_{RA}$ and $\tau_{QoS}$ will be investigated.

ACKNOWLEDGMENTS

The research reported here has been partially supported by the NSF through award 1646229 (CPS-ECCS). A. Coffman is partially supported by University of Florida’s Graduate School Preeminence Award.

REFERENCES

[1] Arian Aghajanzadeh and Peter Therkelsen. 2019. Agricultural demand response for decarbonizing the electricity grid. Journal of Cleaner Production 220 (2019), 827 – 835.
[2] K.J. Arrow. 1958. Studies in the Mathematical Theory of Inventory and Production. Stanford University Press.
[3] Prabir Barooah. 2019. Smart Grid Control: An Overview and Research Opportunities. Springer Verlag. Chapter Virtual energy storage from flexible loads: distributed control with QoS constraints, 99–115.
[4] Prabir Barooah, Ana Bušić, and Sean Meyn. 2015. Spectral decomposition of demand side flexibility for reliable ancillary service in a smart grid. In 46th Hawaii International Conference on System Science (HICSS), invited paper.
[5] Saied Bashash and Hosam K Fathy. 2013. Modeling and control of aggregate air conditioning loads for robust renewable power management. IEEE Transactions on Control Systems Technology 21, 4 (2013), 1318–1327.
[6] Duncan S. Callaway. 2009. Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy. Energy Conversion and Management 50, 5 (2009), 1389 – 1400. https://doi.org/10.1016/j.enconman.2008.12.012
[7] N. J. Cummerdella, R. W. Moyer, Y. Chen, and S. P. Meyn. 2018. An Energy Storage Cost Comparison: Li-ion Batteries vs Distributed Load Control. In 2018 Clemson University Power Systems Conference (PSC), 1–6.
[8] N. Cammerdella, J. Mathias, M. Keiner, A. Bušić, and S. Meyn. 2018. Balancing California’s Grid Without Batteries. In Proceedings of the 2018 Conference on Decision and Control.
[9] Tom C. Chang. 1973. Time interlock for air conditioning compressor and the like. US Patent 3,774,082.
[10] Yue Chen, Md Umar Hashmi, Joel Mathias, Ana Bušić, and Sean Meyn. 2018. Distributed control design for balancing the grid using flexible loads. In Energy Markets and Responsive Grids. Springer, 383–411.
[11] Dingyi Cheng, Wen Zhang, and Kai Wang. 2019. Hierarchical reserve allocation with air conditioning loads considering lock time using Benders decomposition. International Journal of Electrical Power & Energy Systems 110 (2019), 293 – 308. https://doi.org/10.1016/j.ijepes.2019.03.014
[12] Austin Coffman, Ana Bušić, and Prabir Barooah. 2018. Virtual Energy Storage from TCLs using QoS preserving local randomized control. 5th ACM international conference on systems for the built environment 5 (Jan 2018).
[13] Austin R. Coffman, Ana Bušić, and Prabir Barooah. 2018. A Study of Virtual Energy Storage From Thermistically Controlled Loads Under Time-Varying Weather Conditions. In 5th International High Performance Buildings Conference at Purdue. 10.
[14] Austin R. Coffman, Ana Bušić, and Prabir Barooah. 2019. Aggregate capacity for TCLs providing virtual energy storage with cycling constraints. accepted at, IEEE CDC/19.
[15] Michael Grant and Stephen Boyd. 2014. CVX. Matlab Software for Disciplined Convex Programming, version 2.1.
[16] He Hao, Anupama Kowli, Yashen Liu, Prabir Barooah, and Sean Meyn. 2013. Ancillary Service for the Grid Via Control of Commercial Building HVAC Systems.

APPENDIX

We derive the two inventory model (44) and prove Lemma 1 here. The derivation of the inventory model will make use of the indicator function,

$$I_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

Derivation of (44)

We derive here the inventory model, $n_k^on = n_k^on - n_k^off$. We start by identifying that at any given time index $k$ a TCL can do one of four things in regard to its mode state: (i) switch on, (ii) remain on, (iii) switch off, and (iv) remain off. We describe these four possibilities in set form as,

$$X_{s}^{on} = \{m_{k-1}^{j} = 1, m_{k-1}^{i} = 0\}, \quad X_{s}^{off} = \{m_{k}^{j} = 0, m_{k}^{i} = 1\},$$

$$X_{r}^{on} = \{m_{k}^{j} = 1, m_{k}^{i} = 1\}, \quad X_{r}^{off} = \{m_{k-1}^{j} = 0, m_{k-1}^{i} = 1\}.$$ 

For the $j^{th}$ TCL utilizing $z = \{m_{k}^{j}, m_{k-1}^{j}\}$ and $\Delta m_{k} = m_{k} - m_{k-1}$ we write

$$\Delta m_{k} = 1 \cdot (X_{on}^{j}(z) \Delta m) + 1 \cdot (X_{off}^{j}(z) \Delta m).$$

If a TCL remains on or off, then $\Delta m_{k} = 0$ so the above is equivalent to,

$$\Delta m_{k} = 1 \cdot (X_{on}^{j}(z) \Delta m) + 1 \cdot (X_{off}^{j}(z) \Delta m).$$

In American Control Conference. 467–472.
[18] H. Hao, B. M. Sanandaji, K. Poulla, and T. L. Vincent. 2015. Aggregate Flexibility of Thermastically Controlled Loads. IEEE Transactions on Power Systems 30, 1 (Jan 2015), 189–198. https://doi.org/10.1109/TPWRS.2014.2328865
[19] Mingsi Liu, Stef Peeters, Duncan S Callaway, and Bert J Claessens. 2019. Trajectory tracking with an aggregation of domestic hot water heaters: Combining model-based and model-free control in a commercial deployment. IEEE Transactions on Smart Grid (2019).
[20] M. Liu and Y. Shi. 2016. Model Predictive Control of Aggregated Heterogeneous Second-Order Thermistically Controlled Loads for Ancillary Services. IEEE Transactions on Power Systems 31, 3 (May 2016), 1963–1971.
[21] D. Madjidjian, M. Rozeebani, and M. A. Dahleh. 2018. Energy Storage From Aggregate Deferrable Demand: Fundamental Trade-Offs and Scheduling Policies. IEEE Transactions on Power Systems 33, 4 (July 2018), 3573–3586. https://doi.org/10.1109/TPWRS.2017.2766144
[22] Johanna L. Mathieu, Mark Dyson, and Duncan S. Callaway. 2012. Using residential electric loads for fast demand response: The potential resource and revenues, the costs, and policy recommendations. In In Proceedings of the ACEEE Summer Study on Buildings.
[23] Johanna L. Mathieu, Stephan Koch, and Duncan S. Callaway. 2013. State estimation and control of electric loads to manage real-time energy imbalance. IEEE Transactions on Power Systems 28, 1 (2013), 430–440.
[24] Sean Meyn. 2008. Control techniques for complex networks. Cambridge University Press.
[25] B. M. Sanandaji, T. L. Vincent, and K. Poulla. 2016. Ramping Rate Flexibility of Residential HVAC Loads. IEEE Transactions on Sustainable Energy 7, 2 (April 2016), 865–874. https://doi.org/10.1109/TSTE.2015.2497236
[26] Wei Zhang, Jianming Lian, Chin-Yao Chang, and Karanjit Kalsi. 2013. Aggregated modeling and control of air conditioning loads for demand response. IEEE transactions on power systems 28, 4 (2013), 4655–4664.
[27] L. Zhao, W. Zhang, H. Hao, and K. Kalsi. 2017. A Geometric Approach to Aggregate Flexibility Modeling of Thermastically Controlled Loads. IEEE Transactions on Power Systems 32, 6 (Nov 2017), 4721–4731. https://doi.org/10.1109/TPWRS.2017.2674699
[28] C. Zaras, S. You, H. W. Bindner, and E. Vertosick. 2018. A New Method for Handling Lockout Constraints on Controlled TCL Aggregations. In 2018 Power Systems Computation Conference (PSCC). 1–7. https://doi.org/10.23919/PSCC.2018.8442907
Summing the above over the \( j \) index and dividing by \( N \) we obtain,

\[
n_{k}^{on} = n_{k-1}^{on} + s_{k-1}^{on} - s_{k-1}^{off},
\]

which is the desired result. \( \square \)

**Derivation of (26)**

We derive here the inventory model, \( y_{k}^{on} = y_{k-1}^{on} + s_{k-1}^{on} - s_{k-1}^{off} \). For ease, in this derivation we replace \( r_{k}^{on} \) with \( r \). We start by identifying that for a TCL at time index \( k \), there are three possible outcomes regarding its "stuck on" status: (i) it just becomes stuck on, (ii) it just becomes un-stuck on, and (iii) it remains stuck on.

For the \( j^{th} \) TCL we write,

\[
y_{k}^{on,j} = I_{X_{j}^{*}}(z_{1}) y_{k-1}^{on,j} + I_{X_{j}^{+}}(z_{2}) y_{k-1}^{on,j} - m_{k-r}^{j} + \gamma_{k}^{on} + \gamma_{k}^{off}.
\]

Summing the above over the \( j \) index and dividing by \( N \) we obtain,

\[
y_{k}^{on} = \frac{1}{N} \sum_{j=1}^{N} \left( I_{X_{j}^{*}}(z_{1}) + I_{X_{j}^{+}}(z_{2}) \right) y_{k-1}^{on,j} + s_{k-1}^{on} - s_{k-1}^{on},
\]

where \( I_{X_{j}^{*}}(z_{1}) + I_{X_{j}^{+}}(z_{2}) = 1 \) since one event must occur. \( \square \)

**Proof of Lemma 1**

Let \( \otimes \) denote the matrix Kronecker product and \( I_{N_{t}} \), the \( N_{t} \times N_{t} \) identity matrix. The Hessian of \( J(\psi) \) in (61) is,

\[
\frac{\partial^{2} J(\psi)}{\partial \psi^{2}} = I_{N_{t}} \otimes \Xi,
\]

which is positive definite as both \( I_{N_{t}} \) and \( \Xi \) are positive definite.

This proves that the objective function is strictly convex, as the Hessian of \( J(\psi) \) is a positive definite matrix.

To show that the inequality constraints (33) and (43) are feasible it suffices to show that the upper bound of each constraint is always larger than the lower bound. This is trivially satisfied for both constraints as \( s_{k}^{off} + y_{k}^{on} \leq 1 \).

To show that the constraint set \( \Omega = \bigcap_{k=1}^{N_{t}-1} \Omega_{k} \) is non-empty, we start with the baseline trajectory \( r_{k} = 0 \), \( z_{k} = 0 \) and show that this would allow for a \( \bar{\psi}_{k} \in \Omega_{k} \) for all \( k \). If \( r_{k} \) is zero for all \( k \) then \( n_{k}^{on} = n_{k}^{on} \) constant, which due to (45) implies that \( u_{k-1}^{on} = u_{k-1}^{off} \) for all \( k \). Consequently, this implies that \( u_{k-1}^{on} = u_{k-1}^{off} = \bar{u}_{k-1} \) for all \( k \) and that \( y_{k}^{on} = y_{k}^{off} = \bar{y}_{k} \) for all \( k. \) Since the inequality constraints are always feasible, then for all \( k \) and for all \( n_{k}^{on} \) there exists an element \( \bar{\psi}_{k} = [0, 0, \bar{u}_{k}, \bar{y}_{k}, \bar{y}_{k}] \) such that \( \bar{\psi}_{k} \in \Omega_{k}. \) Thus the set \( \Omega \) is non-empty.

To show convexity, we use the fact that the intersection of a finite number of convex sets is convex. Each \( \Omega_{k} \) is convex as the inequality constraints are convex sets and the equality constraints are affine. Thus, \( \Omega \) is convex as it is the finite intersection of convex sets. \( \square \)