Strategic scheduling of residential energy consumers

Adrian Albert[1,*] and Ram Rajagopal[2]

***Abstract*— We propose the *strategic scheduling problem* of an energy utility that administers a large residential population and may ask individual consumers to curtail part of their HVAC usage profile up to a certain *effort budget*, in such a way that the aggregate reductions follow a desired day-ahead goal profile. Each consumer is described by a forecast of their *thermal response profile* computed using a statistical model that decomposes smart meter data into a thermally-sensitive component and an intentional usage component. We propose an algorithm for computing individually-tailored optimal action schedules (e.g., automated Demand-Response control or marketing calls) for thermal energy consumption for a large sample of consumers that is based on solving a convex program. Then, we describe an approximate version of the original scheduling problem that is both interpretable and faster to compute. For this, we recast strategic scheduling as a discrete, set selection problem in which the operator is constrained by what effort structures it can request from the consumers. We propose an efficient algorithm for selecting optimal sets of consumers based on optimizing non-monotone submodular functions.**

I. Introduction

Whose consumption can a utility company influence most cost-effectively? How much effort should it ask for, and at what time? Addressing these questions can aid utilities in creating tailored, targeted marketing actions and operational controls that allow principled shaping of aggregate demand. Here we focus on heating, ventilation, and air conditioning (HVAC), accounting for $\sim 25\%$ of residential electricity use.

Our setup is composed of *i)* a consumer model and *ii)* an operator's problem. The former part assumes a model of individual energy use [1] that interprets observed consumption readings at any point in time as the outcome of consumer decisions of whether to use cooling (air conditioning, AC), heating (furnace), or no temperature-sensitive appliance. The model performs a coarse decomposition of consumption into a temperature-sensitive part - which we call *controllable* or *flexible* - and a temperature-insensitive part - the *baseload*.

On its part, the operator decides whom to send a control signal based on *i)* the statistics of consumption of individual consumers, *ii)* the system-wide aggregate demand profile target, and *iii)* constraints on the effort allowed for each consumer. We call a *schedule* the time sequence of these control signals (e.g., automated Demand-Response events) over the course of the planning horizon. We formulate the operator's problem as a convex program that minimizes the mismatch between the aggregated reduction profile obtained from controlling HVAC consumption and a goal profile. Since in practice tailored control signals for each individual may not be feasible (e.g., due to contracts and policy), we develop an algorithm that almost-optimally assigns a schedule from a fixed set to each consumer.

Related literature

Previous studies on analytic techniques for energy demand management have taken a primarily *descriptive* approach. For example, [2] uses 15-minute resolution smart meter data from ~ 200 customers of an utility company in Germany to cluster consumers according to their daily consumption profiles, and to argue for different pricing schemes for each of the clusters. Similar benchmarks are learned in [3], with an emphasis on describing intra-day usage through a small number of recurring consumption profile patterns. However all these approaches typically ignore the effect of weather and tend to average out the high volatility observed in residential data.

Thermal energy consumption in buildings has long been a topic of high interest in the energy controls literature. Studies have typically employed engineering assumptions and regression models to estimate operational quantities such as thermostat setpoints [4]. More complex algorithms have been proposed to disaggregate specific end-uses (HVAC being a major one) from whole-home energy readings (see [5] for a review); they typically leave out how information about what is being used at every point in time may be useful for energy utilities in operating the grid. The user model that we use here [1] achieves a coarser thermal disaggregation, with the purpose of developing high-level metrics that may serve for DR segmentation and targeting; however it is much less computationally intensive and more interpretable.

There is a large body of literature on normative models for Demand-Response controls that typically assume simple user models that offer closed-form or computationally-inexpensive solutions. E.g., in [6] the authors propose a model of scheduling deferrable loads which is formulated as optimizing a submodular cost function. In [7] simulated profiles of water heating are used to describe control strategies across a user base. Yet ours is to our knowledge the first study that illustrates the use of well-performing user consumption models to the (normative) problem of "smart" HVAC scheduling.

The rest of the paper is structured as follows. In Section II we introduce the strategic scheduling problem. Section IV describes the data used and the inputs to the optimization resulted from the consumer model. Section V describes effort schedules mined from data. Section VI concludes the paper.

II. Problem formulation

We consider an energy utility company (system operator, or SO) that services a population of N energy consumers in a given geography, who experience similar weather patterns. For each consumer, the operator observes hourly usage readings x_t

[1]C3 Energy, Inc. Work done while at Stanford University.
[2]Civil and Environmental Engineering Department, Stanford University.
[*]Correspondence to adrian.albert@c3energy.com
We would like to thank Baosen Zhang and Chin Woo Tan for useful discussions in preparing this paper.

and temperature levels T_t, where t is a time index, $t = 1, \ldots, \mathscr{T}$, typically for a year, so $\mathscr{T} = 365 \times 24 = 8760$. Using the historical, high-resolution smart meter and temperature data $\{x_t, T_t\}_{t=1}^{\mathscr{T}}$, SO may build, at the individual level, rich statistical models that estimate the consumption response to variation in temperature. We denote by $X(t, T_t)$ the consumption at time t and temperature level T_t specified by the model. These models capture much of the variation observed in consumption time series and have direct interpretability. Then, at an initial time t_0 and given a forecast temperature profile $\mathbf{T} \equiv \{T\}_1^\tau \equiv \{T(t_0), \ldots, T(t_0 + \tau - 1)\}$ ($\mathbf{T} \in \mathbb{R}^\tau$, assumed very good) for each consumer $i = 1, \ldots, N$ the operator may forecast the distributions of the non-controllable, base load profile (all end uses but thermal) $b_i(t)$ and of the controllable thermal response rate $a_i(t)$ (the rate of change of overall consumption with temperature) using a dynamic model as in [1]. These quantities are Gaussian random vectors:

$$\text{Thermal response rate } \left[\frac{\text{kWh}}{{}^\circ F}\right] \qquad \mathbf{a}_i \sim \mathscr{N}(\bar{\mathbf{a}}_i, W_i) \qquad (1)$$

$$\text{Base load } [\text{kWh}] \qquad \mathbf{b}_i \sim \mathscr{N}(\bar{\mathbf{b}}_i, V_i), \qquad (2)$$

where $\mathbf{a}_i, \mathbf{b}_i, \bar{\mathbf{a}}_i$, and $\bar{\mathbf{b}}_i \in \mathbb{R}^\tau$, $V_i, W_i \in \mathbb{R}^{\tau \times \tau}$, and $a_i(t) \sim \mathscr{N}(\bar{a}_i(t), w_i^2(t))$, $b_i(t) \sim \mathscr{N}(\bar{b}_i(t), v_i^2(t))$ for $t = t_0, \ldots, t_0 + \tau - 1$. We denote the set of all such thermal response profile forecasts by $\Omega \equiv \{\mathbf{a}_i |\ i = 1, \ldots, N\}$.

A. The system operator's problem

Matching demand and supply variability. The SO wishes to extract demand-side flexibility from the distributed residential consumer portfolio by modifying the HVAC usage of some of these consumers over a planning horizon of $\tau \ll \mathscr{T}$ hours (here next-day, $\tau = 24$ hours) such that this flexibility may be used to stabilize the variability in the generation profile. The example desired goal profile $g(t)$ ($\mathbf{g} \in \mathbb{R}_+^\tau$) over the planning horizon we use in this paper is shown in Figure 1. Here $\mathbf{g}$ is assumed known and deterministic, and may describe e.g., the deviations in supply (from the baseline provided by conventional sources) in a certain region given by renewable sources.

Control schedules. The operator issues requests for *effort schedules* $u_i(t)$ ($\mathbf{u}_i \in \mathbb{R}^\tau$) to control the HVAC usage for certain consumers $i = 1, \ldots, N$. The quantity $u_i(t)$ is the requested number of degrees F that the thermostat setpoint be modified at time t by consumer i. Note that a zero schedule $\mathbf{u}_i = (0, \ldots, 0)$ is equivalent to not requesting participation from consumer i.

As a result of the request $u_i(t)$, the utility receives the energy reductions $\delta_i(t)$ from user i at time t. Then we have for the aggregate reductions profile Δ:

$$\Delta = \sum_{i=1}^{N} \delta_i \text{ with } \Delta \in \mathbb{R}^\tau. \qquad (3)$$

The market setting. Moreover, the SO operates in a market context where it incurs costs for failing to address the mismatch of demand and supply. We assume that deviations from the goal profile $\mathbf{g}$ carry a time-varying penalty $\mathbf{q}$ (in \$/kWh, $\mathbf{q} \in \mathbb{R}^\tau$). Figure 1 presents a possible structure of $q(t)$ based on the time-of-use rate for PG&E in California.

The mismatch between the variability in demand and that in supply is an amount of energy that must be acquired (or sold

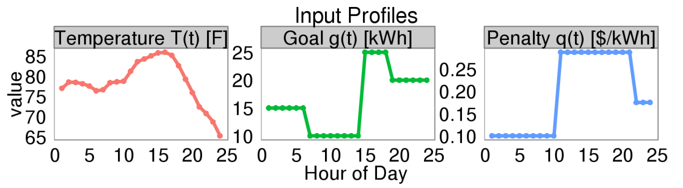

Fig. 1. Strategic scheduling aims to match the goal reductions profile $g(t)$ through aggregate reductions $\Delta(t)$ obtained by controlling thermally-sensitive consumption at the individual level. Deviations are penalized at a rate $q(t)$. Forecasts of thermal response are based on the input temperature profile $T(t)$.

on) the electricity markets. Here, the cost of purchasing additional generation is in general quadratic in the amount desired, as the marginal costs of generation increase approximatively linearly with the generation needed. As such, the SO's cost may be expressed via a quadratic form:

$$C(\{\mathbf{u}_i\}_{i=1}^N) = \|Q^{1/2}(\mathbf{g} - \Delta(\{\mathbf{u}_i\}_{i=1}^N))\|_2^2, \qquad (4)$$

where $Q = \text{diag}(\mathbf{q})$ and we denoted by $\{\mathbf{u}_i\}_{i=1}^N \equiv (\mathbf{u}_1, \ldots, \mathbf{u}_N)$. This assumes that both positive and negative deviations from the goal profile $\mathbf{g}$ are equally undesirable; that is, the SO would like to shape demand reductions such that it best follows $\mathbf{g}$.

The cost of control. However, managing demand may be costly, so not every consumer's usage can be economically adjusted at all times. Each consumer may only be asked to take action (modify their HVAC consumption) up to a certain *effort budget* β that may be either constant or varying across consumers. The SO's task is to minimize the expected cost:

$$\min_{\{\mathbf{u}_i\}} \ \mathbb{E}\ C(\{\mathbf{u}_i\}_{i=1}^N) \qquad (5)$$

$$\text{s.t. } \mathbf{0} \leq \mathbf{u}_i \leq \mathbf{1},\ \forall i = 1, \ldots, N, \qquad (6)$$

$$\mathbf{u}_i^T \mathbf{1} \leq \beta,\ \forall i = 1, \ldots, N, \qquad (7)$$

Note that in the case where there is no budget constraint (7), the objective above is separable, and the minimization may be performed over each time period separately. However, with the addition of the constraint, there appears an *strategic* trade-off between controlling at a given time (which incurs a certain cost and offers a payoff), or retaining the right to control later.

B. The consumer's problem

Based on the input weather profile $\{T_t\}_1^\tau$, the operator may forecast the thermal response (the change in consumption brought about by a change in outside temperature) of individual consumers using a model such as the one developed in [1]. The individual consumer i receives an effort schedule request $u_i(t)$ from the SO; he then may decide how to respond at each time. The SO thus requires a change $u_i(t)\Delta T$ in the thermostat setpoint, where ΔT is the maximum acceptable hourly effort (e.g., $\Delta T = 5^\circ F$), which changes the perceived temperature to $\hat{T}_t = T_t - u_i(t)\Delta T$. This induces a new level of consumption, from $X(t, T_t)$ to $X(t, T_t - u(t)\Delta T)$. Then the change in (thermal) consumption may be written as

$$\begin{aligned} \delta_i(t) \equiv& X_i(t, T_t) - X_i(t, T_t - u_i(t)\Delta T) \\ \approx& a_i(t) u_i(t) \Delta T, \text{ or} \\ \delta_i =& A_i \mathbf{u}_i \text{ in vector form,} \end{aligned}$$

where $A_i = \Delta T \cdot \mathrm{diag}(\mathbf{a}_i)$, and we absorb ΔT into $\mathbf{a}_i$ for simplicity of notation. Since this is a simple affine transformation of a_i, the change in estimated thermal consumption profile will again follow a Gaussian, $\delta_i \sim \mathcal{N}(\bar{\delta}_i, \hat{W}_i)$, where $\bar{\delta}_i(t) = \bar{a}_i(t)u_i(t)$ and $\hat{W}_i = U_i^T W_i U_i$, with $U_i = \mathrm{diag}(\mathbf{u}_i)$. Then it may be easily shown that Δ follows a Gaussian, with $\mathbb{E}[\Delta] = \sum_{i=1}^N \bar{A}_i u_i$ and $\mathrm{Var}[\Delta] = \sum_{i=1}^N U_i W_i U_i^T = \sum_{i=1}^N \hat{W}_i$, where $\bar{A}_i \equiv \mathrm{diag}(\bar{\mathbf{a}}_i)$.

C. Fixed set of schedules

Computing tailored control schedules for each individual consumer may have certain downsides:

1) tweaking consumption in a tailored way for many thousands of consumers may lead to an actual *increase* in volatility on the grid and may adversely affect grid stability;
2) managing thousands of control schedules may be prohibitively complex operationally;
3) reliably estimating $N\tau$ control parameters carries high computational and data requirements for large N.

Thus, we consider the scenario where each consumer may only receive (or choose from) a small set of types of schedules. These schedules may be specified in a contract or in a marketing campaign, in which case they should be simple to convey to the consumer (e.g., "turn up the thermostat setpoint by $3°F$ after 4 pm"). We denote the set of schedules available to consumer i by S_i, and allow it to include both the "null" schedule $\mathbf{0} = (0, ..., 0) \in \mathbb{R}^\tau$ that encodes not selecting user i at all, and simple step effort profile structures. That is,

$$\mathscr{U}_i \equiv \mathbf{0} \cup \{\tilde{\mathbf{u}}_i(\eta, \beta) : \; |\tilde{\mathbf{u}}_i(\eta; \beta)| = \beta\}, \tag{8}$$

where β and η are a pre-set integer parameters, and

$$\tilde{u}_i(t; \eta, \beta) = \begin{cases} 0 & \text{if } t \in [1, \eta - 1] \cup [\eta + \beta + 1, \tau] \\ 1 & \text{if } t \in [\eta, \eta + \beta] \end{cases} \tag{9}$$

We denote $\mathscr{Y} = \cup_{i=1}^N \mathscr{U}_i$. The scheduling problem may then be written as one of selecting a subset $\mathscr{A} \subseteq \Omega$ of consumers, and for each consumer $e \in \mathscr{A}$ a corresponding schedule $\tilde{\mathbf{u}}_e \in \mathscr{U}_e$ with $\mathscr{U} \equiv \{\tilde{\mathbf{u}}_e \in \mathscr{U}_e | e \in \mathscr{A}\}$, such that a match is achieved between the variability in demand and that in supply:

$$\begin{aligned} \min_{\mathscr{A}} \; & \mathbb{E}\; C(\mathscr{U}). \\ \text{s.t. } & \mathscr{U} \equiv \{\tilde{\mathbf{u}}_e \in \mathscr{U}_e | e \in \mathscr{A}\} \end{aligned} \tag{10}$$

This is a discrete optimization of a *submodular* function, as detailed in Appendix VI. Efficient approximate algorithms are known for optimizing such a function (see Section III).

Note that the form (9) considered here has been chosen for simplicity, but imposes significant restrictions on the amount of flexibility that can be extracted from the consumers. The difference in the objective between between the approximate and the tailored solutions will depend on how flexible the sets S_i of fixed schedules are that available to each customer i.

III. Computing effort schedules

A. Individually-tailored schedules

The objective in the operator's problem (5) may be recast to a standard Quadratic Programming (QP) formulation. For

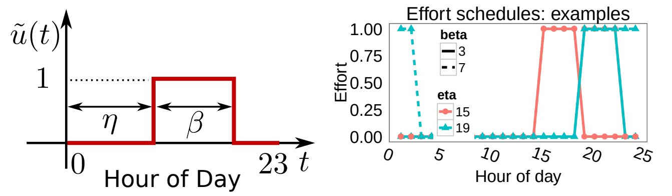

Fig. 2. Types of fixed schedules considered for sets $\mathscr{U}_i$. *Left:* "square wave" schedule set-up; *Right:* example schedules for different values of η and γ.

that, we define $\mathbf{u} \equiv (u_1^T, ..., u_N^T)^T$, with $\mathbf{u} \in \mathbb{R}^{N\tau \times 1}$ and $\mathbf{A} \equiv (A_1, ..., A_N)$, with $\mathbf{A} \in \mathbb{R}^{\tau \times N\tau}$. Then with $\Delta = \mathbf{Au}$, (5) becomes:

$$\begin{aligned} \mathbb{E}[C] &= \mathbb{E}[\Delta^T Q \Delta] - 2\mathbf{g}^T Q \mathbb{E}[\Delta] + \mathbf{g}^T Q \mathbf{g} \\ &= \mathbf{u}^T \mathbb{E}[\mathbf{A}^T Q \mathbf{A}]\mathbf{u} - 2\mathbf{g}^T \bar{\mathbf{A}}\mathbf{u} + \mathbf{g}^T Q \mathbf{g} \\ &= \mathbf{u}^T \mathbf{H} \mathbf{u} - 2\mathbf{h}^T \mathbf{u} + \mathbf{c}, \end{aligned} \tag{11}$$

where the block matrix $\mathbf{H}$ may be constructed as

$$\mathbf{H}_{i,j} = \begin{cases} \bar{A}_i^2 Q + Q W_i & \text{if } i = j \\ \bar{A}_i^T Q \bar{A}_j & \text{if } i \neq j \end{cases} \tag{12}$$

With $\mathbf{1}_N$ and $\mathbf{1}_{\tau \times \tau}$ representing, in turn, a $N \times 1$ vector and a $\tau \times \tau$ matrix with all entries of 1, we further define

$$\mathbf{G} = \mathbf{1}_{\tau \times \tau} \otimes \mathrm{diag}(\mathbf{1}_N). \tag{13}$$

Then algorithm 1 outlines the procedure for computing the tailored effort schedules for each individual consumer.

Algorithm 1 Computing tailored individual schedules.

Require: Thermal response profiles $\{a_i\}_{i=1}^N$; goal profile $\mathbf{g}$; effort budget β.
Ensure: Effort schedules $\{u_i\}_{i=1}^N$.

1) Compute $\mathbf{H}$ using Equation (12), $\mathbf{h}$ and $\mathbf{c}$ from Equation (11), and $\mathbf{G}$ using Equation (13).
2) Solve QP

$$\begin{aligned} \min_{\mathbf{u}} \; & \mathbf{u}^T \mathbf{H} \mathbf{u} - 2\mathbf{h}^T \mathbf{u} + \mathbf{c} \\ \text{s.t. } & \mathbf{G}\mathbf{u} \leq \beta \\ & \mathbf{0} \leq \mathbf{u} \leq \mathbf{1}, \end{aligned}$$

using a high-performance solver, e.g., `MOSEK` via the `Rmosek` interface to the statistical computing language `R`.

B. Fixed sets of schedules

We now consider selecting a subset $\mathscr{A} \subseteq \Omega$ whereby consumer i is offered a fixed schedule $\tilde{\mathbf{u}}_i \in \mathscr{U}_i$, where $\mathscr{U}_i$ is fixed in advance. The schedules are generated as "square waves", i.e., a brief period of constant effort β/γ of duration γ, starting at time $1 \leq \eta \leq 24$, as depicted in Figure 2 (left panel). Examples of such schedules are given in the right panel of the figure. For our analysis we take all sets $\mathscr{U}_i$ to be identical; however depending on the context it may be advantageous to consider different $\mathscr{U}_i$'s for different consumers, e.g., to reflect consumer preferences or particular individual contract terms. Minimizing $\mathbb{E}[C(\mathscr{A})]$ is equivalent to maximizing $f(\mathscr{A}) = \mathbb{E}[C(\varnothing)] - \mathbb{E}[C(\mathscr{A})]$, which is a submodular, non-monotonic function (see Appendix VI), for which efficient, approximate algorithms exist [8], [9], [6]. Algorithm 2 presents one such greedy approach with an approximation factor of $\alpha = \frac{1}{3} - \frac{\varepsilon}{n}$, with $\varepsilon > 0$ small and $n = |\mathscr{A}|$.

Algorithm 2 Greedy maximization of submodular functions.

Require: set $\Omega = \{a_i\}_{i=1}^{N}$; Sets $\mathscr{U}_i$, $i = 1, \ldots, N$; parameter ε.
Ensure: selected consumer set $\mathscr{A}$, effort schedule set $\mathscr{U}$.

Initialize

1) $e,\ \tilde{\mathbf{u}}_e := \underset{\tilde{\mathbf{u}}_e \in \mathscr{Y}: e \in \Omega}{\operatorname{argmax}}\ f(\mathscr{U})$;
2) $\mathscr{A} := \{e\}$, $\mathscr{U} := \{\tilde{\mathbf{u}}_e\}$, $\mathscr{Y}' := \mathscr{Y} \setminus \mathscr{U}_e$.

Repeat

1) *Step forward:* Find $e \in \Omega \setminus \mathscr{A}$ and $\tilde{\mathbf{u}}_e \in \mathscr{Y}'$ for which

$$f(\mathscr{U} \cup \tilde{\mathbf{u}}_e) > \left(1 + \frac{\varepsilon}{n^2}\right) f(\mathscr{U})$$

`If` $e \neq \varnothing$:

- let $\mathscr{A} := \mathscr{A} \cup \{e\}$, $\mathscr{U} := \mathscr{U} \cup \{\tilde{\mathbf{u}}_e\}$, $\mathscr{Y}' := \mathscr{Y}' \setminus \mathscr{U}_e$;

`Otherwise:`

2) *Step backwards:* Find $e \in \mathscr{A}$ and $\tilde{\mathbf{u}}_e \in \mathscr{U}$ for which

$$f(\mathscr{U} \setminus \tilde{\mathbf{u}}_e) > \left(1 + \frac{\varepsilon}{n^2}\right) f(\mathscr{U})$$

`If` $e \neq \varnothing$:

- $\mathscr{A} := \mathscr{A} \setminus \{e\}$, $\mathscr{U} := \mathscr{U} \setminus \{\tilde{\mathbf{u}}_e\}$, and $\mathscr{Y}' := \mathscr{Y}' \cup \mathscr{U}_e$;

`Otherwise: terminate.`

until `terminate` or $\mathscr{Y}' = \varnothing$.

IV. Experimental setup

In this paper we use real smart meter consumption time series and corresponding weather time series (at the 5-digit zipcode level) for each premise for a sample of 1,493 premises in a hot region around Bakersfield, CA. This is whole-premise data (no individually-monitored appliance data is available) at an hourly level and spans one year from August 30^{th}, 2010 to July 31^{st}, 2011. We chose this region since it is a hot, arid climate where temperatures are relatively consistent throughout the year; as such many consumers here are expected to use Air Conditioning (AC). As the targeting and control strategies discussed here are to be applied in a localized way, it is appropriate to illustrate it on a sample that is indicative of the size of a typical substation. For each consumer in the selected sample we learn a consumption model as in [1] that we then use to make forecasts of thermal response profiles. Figure 3 presents the temperature response and baseload profiles $\mathbf{a} \to (\bar{\mathbf{a}}, W)$ and $\mathbf{b} \to (\bar{\mathbf{b}}, V)$ for each of three example consumers, "Robert", "Joe", and "Lewis". Robert does not have a high thermal DR potential, whereas Lewis might be able to offer more flexibility - yet with greater uncertainty.

V. Results

A. Individually-tailored schedules

We compute individually-tailored schedules using Algorithm 1 for a value of $\beta = 3$. This very flexible (and computationally-intensive) procedure yields an aggregate reductions profile Δ that matches very closely the desired goal

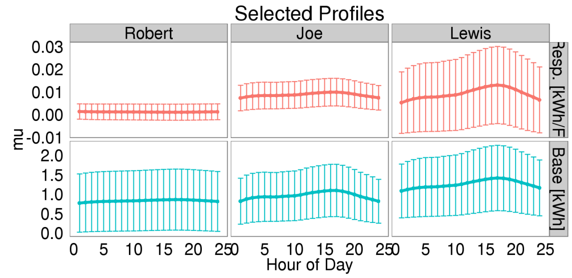

Fig. 3. Forecast thermal response **a** and baseload **b** profiles for three example consumers. The solid lines represent the distribution means, while the error bars represent the diagonal entries in the covariance matrices of the profiles.

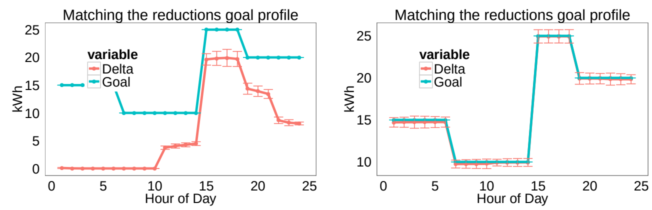

Fig. 4. Match between the goal **g** and aggregate reductions Δ for effort budget $\beta = 1, 2, 3$ (from left to right). Error bars indicate the variance in the aggregate profile around the mean plotted in a solid line. With more flexibility asked of each consumer, a better match is achieved between **g** and Δ.

profile **g**, with 408 out of 1,493 consumers selected (i.e., for which the schedules $\|\mathbf{u}\| > 0$) for the program. Across the population, the most effort is required in the late afternoon and evening hours - when most consumers turn on their AC, while the least effort is requested late morning to early afternoon.

In Figure 4 we exemplify the result of the control strategy on achieving the goal reductions profile **g** for the California time-of-use penalty structure **q**. The expected reductions profile Δ is illustrated for two levels of the effort budget, $\beta_1 = 1$ (left), and $\beta_2 = 3$ (right). Note that with a higher effort burden on the consumers the achieved reductions Δ follow the goal **g** much more closely. Moreover, for low β values (when there is not enough room for flexibility) we observe the "strategic" behavior of the control schedules: the match is better for afternoon hours than early morning, indicating that control is foregone to late afternoon hours when both the cost of acquiring additional generation $q(t)$ is higher and the potential for thermal flexibility observed empirically among the consumers is higher as well. One immediate observation is that for the chosen goal profile **g** (which would be appropriate for local solar energy generation), on a typical end-of-summer day in a hot area such as Bakersfield, an effort budget of $\beta = 3 \times 5 = 15°F$ is enough to extract the required demand-side flexibility. Note that these calculations did not take into account any measure of compliance to control requests; they simply illustrate a "best-case" scenario where compliance is of 100%. The corresponding schedules asked of our example consumers Joe, Robert, and Lewis are illustrated in Figure 5. For low allowed effort ($\beta = 1$), requests for control are mostly in the afternoon, whereas for larger β the same users receive requests that are more spread-out.

Finally, we performed a more extended scenario analysis, in particular being interested in the trade-off between expending resources for enrolling consumers and the averted penalty

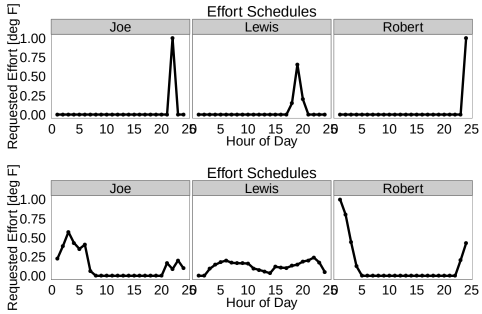

Fig. 5. Effort schedules requested of the example consumers Joe, Robert, and Lewis, for increasing levels of the allowed budget $\beta = 1, 2, 3 \times 5°F$.

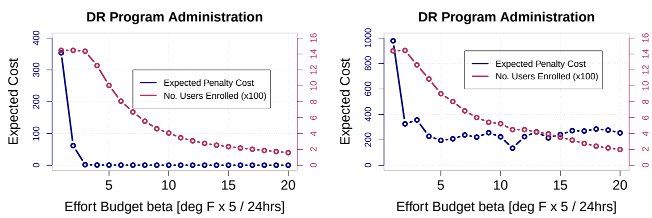

Fig. 6. The trade-off between the effort budget β and goal profile mismatch costs. *Left:* "tailored" solution using individually-tailored schedules; *Right:* approximate solution using fixed schedules.

for missing the reductions goal. In Figure 6 (left panel) we illustrate the dependence of $\mathbb{E}[C]$ and total number of enrolled consumers with the effort budget β. As expected, the more flexibility requested from the consumers, the less costly it is to achieve the reductions goal. In particular, the match is close to perfect for $\beta \geq 3 \times 5°F$, which suggests the existence of a "critical budget value" which acts as a lower bound for the amount of effort asked of each consumer in order to achieve the operational goal. The number of consumers needed to enroll falls with β as well - if more effort may be extracted of a single consumer, fewer participants will be needed.

B. Schedule set selection

The second strategy for computing effort schedules that we investigate here is the problem of choosing from a given set of fixed schedules for each consumer. As argued above, this problem is a combinatorial problem for which only an approximate solution may be found in polynomial time; therefore it is of interest to understand how "far" the approximate solution is from the tailored solution described earlier in this paper (when computing tailored schedules for each individual consumer). In particular, for the administration of a DR program focused on enrolling the right consumers and requesting the appropriate effort schedules two quantities are of relevance: the number n of consumers needed (the 'size' of a program) and the expected penalty achieved under a program of a given size.

For the approximate customer set selection problem we defined fixed sets of schedules $\mathscr{U}_i = \{\tilde{\mathbf{u}}(\eta, \beta) | \eta = 2, 4, 6, ..., 20\}$ as introduced in Section II. As above, we repeat the analysis for different values of $\beta = 1, ..., 20$. In Figure 6 (bottom panel) we illustrate the penalty cost (in \$) as a function of the effort budget β. As expected, the approximate solution using much less flexible schedules (but of the same effort budget β) will result in higher penalty costs than the "tailored" solution

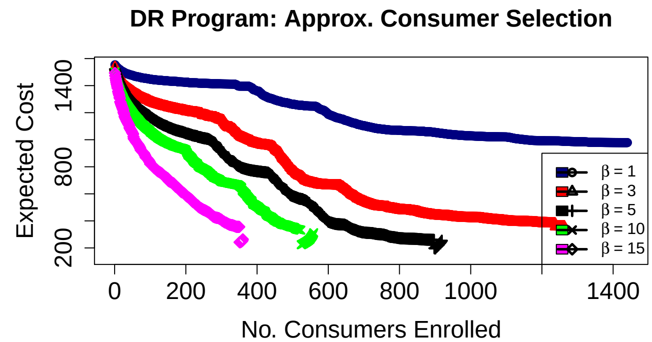

Fig. 7. Penalty when for different values of the effort budget β.

(which for $\beta \geq 3$ has a penalty cost of essentially \$0). We observe the same sharp drop-off in the penalty cost for $\beta \geq 3$, albeit for the approximate solution this cost levels off around \$300 for larger effort budget values (as opposed to \$0 for the tailored solution). This suggests that, when the program is constrained to request only certain fixed, pre-determined effort schedules (e.g., as stipulated in a contract), asking individuals for more effort (in this case controlling the thermostat for longer periods) does not necessarily lead to better program results. Similar to the tailored solution case, the number of enrolled customers will fall, albeit less abruptly, from ~ 1400 (most consumers) for $\beta = 1$ to ~ 200 for $\beta = 20$. Clearly, the tailored value of the "critical" effort budget will vary heavily with the goal $\mathbf{g}$, the penalty structure $\mathbf{q}$, and with the tailored composition of the customer profiles within a region. However, this observation highlights the need for effective use of demand-response flexibility while being mindful to the effort imposed on the customers.

In certain cases however a tailored solution is not available for comparison. As such, it may be of interest to understand "when to stop" with selecting consumers - given that the benefits from keeping to add consumers to the program will be diminishing with the size of the program. We illustrate this experiment in Figure 7 for several selected values of β. With increasing β the expected penalty cost curves exhibit increasingly larger marginal returns. The curves in the figure may serve as guideline for a desired size of a demand-response program by offering a quantitative assessment of how much expected financial penalty will be incurred by selecting sets of consumers of certain sizes n. This may be relevant in particular in the case (not discussed here) when selecting additional consumers is costly in itself. For example, if a number of $n = 200$ consumers is desired in the program, asking each consumer for a total effort of $\beta = 5 \times 5°F$ (over $\tau = 24$ hours) will result in an expected penalty cost of $\sim \$1,050$, whereas when the effort budget is smaller at $\beta = 3 \times 5°F$, the expected penalty increases to $\sim \$1,200$.

VI. Conclusions

Reducing consumption at certain times of the day is beneficial to the utility financially and is environmentally-friendly since it may avert ramping up of generation capacity. Here we provided a framework for analyzing the capacity for day-ahead demand-side flexibility. We illustrated how detailed, dynamic consumption models of individual household energy consumption may be used in an operational setting to derive optimal actions that a system operator may take to achieve a beneficial aggregate consumption profile. Our findings reinforce two important aspects of running DSM programs:

1) With full flexibility for requests for effort (in terms of amount and timing) there is no need for excessive control effort required of a given individual, if that effort is requested at the right time and from a large pool of consumers;
2) When the flexibility for control requests is constrained (as we have done here with constant effort functions), a significant suboptimality gap occurs. This departure from the optimal solution decreases with the number of consumers available for control, but plateaus at a level given by how flexible control is for each individual consumer. A question we have left for future research is how to best design the functional form of the fixed control schedules to achieve a balance between simplicity and the gap between the tailored and approximate solutions.

APPENDIX

Submodular functions. *Submodularity* is the natural analogue of a convexity on discrete domains [9]. A function $f:\Omega\to\mathbb{R}$ (with Ω discrete) is called *submodular* if

$$f(\mathscr{A}\cup\{e\})-f(\mathscr{A})\geq f(\mathscr{B}\cup\{e\})-f(\mathscr{B}), \tag{14}$$

where $\mathscr{A}\subseteq\mathscr{B}\subseteq\Omega$ and $e\in\Omega\setminus\mathscr{B}$. I.e., adding element e to the smaller set $\mathscr{A}$ results in a higher payoff than adding it to the larger set $\mathscr{B}$. The marginal payoff of adding e to $\mathscr{A}$ is:

$$\rho_e(\mathscr{A})=f(\mathscr{A}\cup\{e\})-f(\mathscr{A}), \tag{15}$$

which according to the definition above will be decreasing in $|\mathscr{A}|$; that is, the function f will have diminishing returns.
Cost function is submodular. Following [6] we re-write the objective function in (10) as

$$\begin{aligned} f(\mathscr{A}) &= \mathbb{E}\left[C(\varnothing)-C(\mathscr{A})\right] && (16)\\ &= \|Q^{1/2}\mathbf{g}\|_2^2-\mathbb{E}\|Q^{1/2}(\mathbf{g}-\Delta(\mathscr{A})\|_2^2 \\ &= \sum_{t=1}^{\tau} q(t)f_t(\mathscr{A}), && (17)\end{aligned}$$

which is a linear combination of functions of the form $f_t(\mathscr{A})=g^2(t)-\mathbb{E}\left[(\Delta(t;\mathscr{A})-g(t))^2\right]$. Note that $f(\varnothing)=0$.

Consider the marginal return of adding element e to $\mathscr{A}$:

$$\begin{aligned}\rho_e(\mathscr{A};t) &= f_t(\mathscr{A}\cup\{e\})-f_t(\mathscr{A})\\ &= \mathbb{E}\left[(\Delta(t;\mathscr{A})-g(t))^2\right]-\mathbb{E}\left[(\Delta(t;\mathscr{A}\cup\{e\})-g(t))^2\right]\\ &= u_e(t)\mathbb{E}\left[a_e(t)(2g(t)-2\Delta(t;\mathscr{A})+a_e(t)u_e(t))\right].\end{aligned}$$

Since all quantities involved are positive and $\mathscr{A}\subseteq\mathscr{B}$:

$$\begin{aligned}\rho_e(t;\mathscr{A})-\rho_e(t;\mathscr{B}) &\propto \mathbb{E}\left[a_e(t)(\Delta(t;\mathscr{B})-\Delta(t;\mathscr{A}))\right]\\ &= \mathbb{E}\left[a_e(t)\sum_{j\in\mathscr{B}\setminus\mathscr{A}}\Delta(t;\mathscr{B})-\Delta(t;\mathscr{A})\right]\\ &\geq 0, && (18)\end{aligned}$$

Then, since f is a linear combination of the submodular functions f_t's (with weights $q(t)$), f is submodular[9].

Computing the objective function. Calculating the expected cost $\mathbb{E}[C(\mathscr{U}]$ may at first appear to be a relatively computationally-intensive task, as it involves forming the matrices in Equation (12). However that is not needed if only the computation of $\rho_u(\mathscr{U})$ is desired. After some simple algebraic manipulations it can be shown:

$$\begin{aligned}\rho_u(\mathscr{U}) &= \mathbb{E}[C(\mathscr{U})]-\mathbb{E}[C(\mathscr{U}\cup\{u\})]\\ &= \sum_{t=1}^{\tau} q(t)\left[2u(t)a(t)\sum_{i=1}^{n}u_i(t)a_i(t)\right.\\ &\quad\left.+u^2(t)(\mathrm{diag}(w)(t)+a^2(t))-2g(t)a(t)u(t)\right]. && (19)\end{aligned}$$

Above, $a\in\Omega$ is the thermal profile corresponding to the schedule u. A similar update scheme may be defined for the operation of eliminating an element from a given set (the backwards step in Algorithm 2). This iterative scheme for computing the cost function results in an efficient algorithm.
Optimizing submodular functions. The problem (10) is NP-hard [9], and as such an exact solution can be found only in exponential time. However a large body of work has been dedicated to approximation algorithms that achieve "good" performance in polynomial time. In particular, greedy algorithms have been studied extensively in the context of maximizing monotone submodular functions, for which famously an approximation factor of $\alpha=1-\frac{1}{e}$ has been shown [10] and good practical performance has been observed.

The objective in (10) used here is, however, not monotone, due to the squared term. As such, we use Algorithm 2 in Section III, which has been shown [8] to achieve an approximation factor of $\alpha=\frac{1}{3}-\frac{\varepsilon}{n}$ with each step in which elements are added or removed from the solution that increase the function value by a multiplicative factor of at least $1+\frac{\varepsilon}{n^2}$, where $n=|\mathscr{A}|$.

REFERENCES

[1] A Albert and R. Rajagopal. Thermal profiling of residential energy use. *Power Systems, IEEE Transactions on*, PP(99):1–10, 2014.
[2] Christoph Flath, David Nicolay, Tobias Conte, Clemens van Dinther, and Lilia Filipova-Neumann. Cluster analysis of smart metering data - an implementation in practice. *Business & Information Systems Engineering*, 4(1), 2012.
[3] Jungsuk Kwac, June Flora, and Ram Rajagopal. Household energy consumption lifestyle segmentation using hourly data. *Forthcoming*, 2013.
[4] Benjamin J. Birt, Guy R. Newsham, Ian Beausoleil-Morrison, Marianne M. Armstrong, Neil Saldanha, and Ian H. Rowlands. Disaggregating categories of electrical energy end-use from whole-house hourly data. *Energy and Buildings*, 50(0):93 – 102, 2012.
[5] C. K. Carmel, G. Shrimali, and A. Albert. Disaggregation: the holy grail of energy efficiency? *Energy Policy*, 2013.
[6] G. O'Brien and R. Rajagopal. A method for automatically scheduling notified deferrable loads. In *American Control Conference (ACC), 2013*, pages 5080–5085, June 2013.
[7] Ning Lu and Yu Zhang. Design considerations of a centralized load controller using thermostatically controlled appliances for continuous regulation reserves. *Smart Grid, IEEE Transactions on*, 4(2):914–921, 2013.
[8] Uriel Feige, Vahab S. Mirrokni, and Jan Vondrák. Maximizing nonmonotone submodular functions. *SIAM J. Comput.*, 40(4):1133–1153, July 2011.
[9] Andreas Krause and Daniel Golovin. Submodular function maximization. In *Tractability: Practical Approaches to Hard Problems (to appear)*. Cambridge University Press, February 2014.
[10] G.L. Nemhauser, L.A. Wolsey, and M.L. Fisher. An analysis of approximations for maximizing submodular set functions–i. *Mathematical Programming*, 14(1):265–294, 1978.