Behavior Modification and Utility Learning via Energy Disaggregation

Lillian J. Ratliff * **Roy Dong** * **Henrik Ohlsson** *
S. Shankar Sastry *

* *University of California, Berkeley, Berkeley, CA 94720 USA*
{ratliffl, roydong, ohlsson, sastry}@eecs.berkeley.edu

Abstract: The utility company has many motivations for modifying energy consumption patterns of consumers such as revenue decoupling and demand response programs. We model the utility company–consumer interaction as a principal–agent problem. We present an iterative algorithm for designing incentives while estimating the consumer's utility function. Incentives are designed using the aggregated as well as the disaggregated (device level) consumption data. We simulate the iterative control (incentive design) and estimation (utility learning and disaggregation) process for examples including the design of incentives based on the aggregate consumption data as well as the disaggregated consumption data.

Keywords: Game Theory, Economic Design, Energy Management Systems

1. INTRODUCTION

Currently, most electricity distribution systems only provide aggregate power consumption feedback to consumers, in the form of a energy bill. Studies have shown that providing device-level feedback on power consumption patterns to energy users can modify behavior and improve energy efficiency (Creyts et al., 2007; Gardner and Stern, 2008; Laitner et al., 2009; Perez-Lombard et al., 2008).

However, the current infrastructure only has sensors to measure the aggregated power consumption signal for a household. Even advanced metering infrastructures currently being deployed have the same restriction, albeit at high resolution and frequency (Armel et al., 2013). Additionally, deploying plug–level sensors would require entering households to install these devices. Methods requiring plug–level sensors are often referred to as *intrusive load monitoring*, and the network infrastructure required to transmit high resolution, high frequency data for several devices per household would be very costly.

A low cost alternative to the deployment of a large number of sensors is *non–intrusive load monitoring*. We consider the problem of nonintrusive load monitoring, which, in the scope of this paper, refers to recovering the power consumption signals of individual devices from the aggregate power consumption signal available to our sensors. This is also sometimes referred to as *energy disaggregation*, and we will use the two terms interchangeably. This problem has been an active topic of research lately. Some works include Dong et al. (2013a); Froehlich et al. (2011); Gupta et al. (2010); Johnson and Willsky (2012); Leeb et al. (1995).

We propose that the utility company should use incentives to motivate a change in the energy consumption of consumers. We assume the utility company cares about the satisfaction of its consumers as well as altering consumption patterns, but it may not be able to directly observe the consumption patterns of individual devices or a consumer's satisfaction function.

In brief, the problem of behavior modification in energy consumption can be understood as follows. The utility company provides incentives to energy consumers, who seek to maximize their own utility by selecting energy consumption patterns. This can be thought of as a control problem for the utility company. Additionally, the utility company does not directly observe the energy consumption patterns of individual devices, and seeks to recover it from an aggregate signal using energy disaggregation. This can be thought of as an estimation problem. Further, the consumer does not report any measure of its satisfaction directly to the utility. Thus, it must be estimated as well.

There are many motivations for changing energy consumption patterns of users. Many regions are beginning to implement revenue decoupling policies, whereby utility companies are economically motivated to decrease energy consumption (Eom, 2008). Additionally, the cost of producing energy depends on many variables, and being able to control demand can help alleviate the costs of inaccurate load forecasting. Demand response programs achieve this by controlling a portion of the demand at both peak and off-peak hours (Mathieu et al., 2012). We propose a model for how utility companies would design incentives to induce the desired consumer behavior.

In this paper, we consider three cases of incentive design. First, we consider the case where the utility company

⋆ The work presented is supported by the NSF Graduate Research Fellowship under grant DGE 1106400, NSF CPS:Large:ActionWebs award number 0931843, TRUST (Team for Research in Ubiquitous Secure Technology) which receives support from NSF (award number CCF-0424422), and FORCES (Foundations Of Resilient CybEr-physical Systems), the European Research Council under the advanced grant LEARN, contract 267381, a postdoctoral grant from the Sweden-America Foundation, donated by ASEA's Fellowship Fund, and by a postdoctoral grant from the Swedish Research Council.

designs an incentive based entirely on the aggregate power consumption signal. Our model also supposes the utility company cares about the consumer's satisfaction, which is an unknown function. Thus, there is an information asymmetry between the utility company and the energy consumer, and the utility company does not know its own cost function. We propose an algorithm to estimate the satisfaction function of the consumer based on the consumer's aggregated power consumption signals in Section 3.

In Section 4.1, we consider the case where the utility company knows the power consumption signal of individual devices and an unknown satisfaction function. In this case, the utility company designs incentives for individual devices, and estimates satisfaction functions for individual devices. This provides a more realistic model for the human consumers. In Section 4.2 we consider the case when the utility company only has access to the aggregated power consumption signal, and uses an energy disaggregation algorithm to recover the power consumption of individual devices. This disaggregated signal is used to allocate incentives, but the results will depend on the accuracy of our estimator, the energy disaggregation algorithm. In Section 5 we simulate two examples of designing incentives while estimating the consumer's satisfaction function; one using the aggregated signal and the other using a disaggregated signal with an error bound. Finally, in Section 6 we make concluding remarks and discuss future research directions.

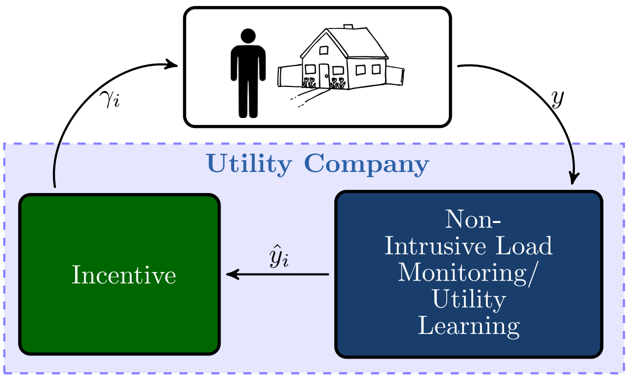

Fig. 1. Closing the Loop: Behavior modification via incentives γ_i abstractly is a *control* problem. The agent decides when to use devices resulting in device level consumption y_i. Non–intrusive load monitoring is abstractly *estimation* in that aggregate power consumption y is observed and disaggregation algorithms are used to estimate device level usage $\hat{y}_i$ for each device. Similarly, utility learning is an estimation problem. The utility company solves the control and estimation problem.

2. INCENTIVE DESIGN PRELIMINARIES

2.1 Principal–Agent Model

A principal-agent problem occurs when the principal interacts with the agent to perform a task, but the agent is not incentivized to act in the principals best interests. This conflict is often the result of asymmetric information between the principal and the agent or a disconnect between their goals and objectives.

2.2 Incentive Design Problem

The principal–agent problem we consider is a leader–follower type problem in which the principal is the leader and the agent is the follower (Laffont and Martimort, 2009). Specifically it is a Stackelberg game (Basar et al., 1995). Both the principal and the agent wish to maximize their pay–off determined by the functions $J_p(v, y)$ and $J_a(v, y)$ respectively. The principal's decision is denoted v; the agent's decision, y; and the incentive, $\gamma : y \mapsto v$. The basic approach to solving the Stackelberg game is as follows. Let v and y take values in $V \subset \mathbb{R}^{n_p}$ and $Y \subset \mathbb{R}^{n_a}$, respectively; $J_p : \mathbb{R}^{n_p} \times \mathbb{R}^{n_a} \to \mathbb{R}$; $J_a : \mathbb{R}^{n_p} \times \mathbb{R}^{n_a} \to \mathbb{R}$. We define the desired choice for the principal as

$$(v^d, y^d) = \arg\max_{v,y} J_p(v, y) \tag{1}$$

subject to $v \in \mathcal{I}_v$ and $y \in \mathcal{I}_y$ where $\mathcal{I}_v$ and $\mathcal{I}_y$ are constraints on v and y respectively. The incentive problem can be stated as follows:

Problem 1. Find $\gamma : Y \to V$, $\gamma \in \Gamma$ such that

$$\arg\max_{y} J_a(\gamma(y), y) = y^d \tag{2}$$

$$\gamma(y^d) = v^d \tag{3}$$

where Γ is the set of admissible incentive mechanisms.

3. INCENTIVE DESIGN USING AGGREGATE POWER SIGNAL

We cast the utility–consumer interaction model in the framework of a principal–agent model in which the utility company is the principal and the consumer is the agent (see Figure 1). The principal's true utility is assumed to be given by

$$J_p(v, y) = g(y) - v + \beta f(y) \tag{4}$$

where $g(\cdot)$ is a concave function of the consumer's energy usage y over a billing period, v is the value of the incentive paid to the agent, $f : Y \to \mathbb{R}$ is the agent's satisfaction function for energy consumption which we assume is concave and β is a multiplying factor capturing the degree of benevolence of the principal. The value of v is constrained to be greater than or equal to zero, i.e. the principal should not take additional money away from the consumer on top of the cost of their usage, and less than some maximal amount the principal is willing to pay to the consumer $v^{\max}$. Let us denote the constraint on v with $\mathcal{I}_v = [0, v^{\max}]$. Similarly, let $y \in \mathcal{I}_y = [0, y^{\max}]$. In a regulated market with revenue decoupling in place, a simplified model may consider

$$g(y) = -y \tag{5}$$

representing the fact that the utility wants the agent to use less energy. Similarly, if the utility company has aspirations to institute a demand response program, a simplified model may consider

$$g(y) = -\left(y - y^{\text{ref}}\right)^2 \tag{6}$$

where y^{ref} is the reference signal prescribed by the demand response program.

The agent's true utility is assumed to be

$$J_a(\gamma(y), y) = -py + \gamma(y) + f(y) \tag{7}$$

where p is the price of energy set and known to both the agent and the principal and $\gamma : Y \to \mathbb{R}$ is the incentive mechanism. Incentives are designed by solving Problem 1. Let Γ be the set of concave functions from Y to $\mathbb{R}$. Throughout the paper, for the sake of analysis, we restrict γ to live in Γ. Under the assumption that f is concave and $\gamma \in \Gamma$, then J_a is concave. The principal does not know the agent's satisfaction function $f(\cdot)$, and hence, must estimate it as he solves the incentive design problem. We will use the notation $\hat{f}$ for the estimate of the satisfaction and $\hat{J}_p$ and $\hat{J}_a$ for the player's cost functions using the estimate of f.

We propose an algorithm for iteratively estimating the agent's satisfaction function and choosing the incentive $\gamma(\cdot)$. We do so by using a polynomial estimate of the agent's satisfaction function at each iteration. Suppose that $\gamma^{(0)}$ and $\gamma^{(1)}$ are given a priori. At each iteration the principal issues an incentive and observes the agent's reaction. The principal then uses the observations up to the current time along with his knowledge of the incentives he issued to estimate the agent's utility function. Formally, at the k–th iterate the principal will observe the agent's reaction $y^{(k)}$ to a delivered incentive $\gamma^{(k)}$. The agent's reaction $y^{(k)}$ is optimal with respect to

$$J_a(\gamma^{(k)}(y), y) \tag{8}$$

in that $y^{(k)}$ maximizes the true utility J_a under the incentive $\gamma^{(k)}$. We use the observations $y^{(0)}, \dots, y^{(k)}$ to estimate the parameters in the agent's satisfaction function given by

$$\hat{f}^{(k)}(y) = \sum_{i=0}^{j} \alpha_i y^{i+1} \tag{9}$$

where j is the order of the polynomial estimate to be determined in the algorithm. We do not need a constant term in our estimate because it does not affect the optimization problem. We assume that $y^{(i)} \in (0, y^{\max})$; otherwise, we terminate the algorithm. We know that a necessary condition for a global minimum when

$$J_a(\gamma(y), y) = -py + \gamma(y) + \hat{f}(y) \tag{10}$$

is concave and Y is open is that

$$\nabla J_a(\gamma(y^*), y^*) = 0 \tag{11}$$

(Bertsekas, 1999). Similarly, Equation (11) is necessary for a local minimum of $\hat{J}_a \in C^1(Y)$ with Y open. The principal only knows his estimate of J_a; hence, our algorithm prescribes that he uses his belief about J_a, namely $\hat{J}_a$ in the necessary condition (11). In the case that $\hat{f}$ is concave and under our assumption that the agent is rational and hence plays optimally, the observation $y^{(i)}$ is a global optimum at iteration i. In the case when $\hat{f}$ is not concave, the agent plays myopically; we can only guarantee that the observation $y^{(i)}$ is a local optimum. In both cases, we can use the necessary condition

$$\nabla \hat{J}_a^{(i)}(\gamma(y^{(i)}), y^{(i)}) = 0 \tag{12}$$

for each of the past iterates $i \in \{0, \dots, k\}$ to determine estimates of the coefficients in $\hat{f}^{(k)}$. At the k–th iteration, we have measurements $y^{(0)}, \dots, y^{(k)}$ and we know $\gamma^{(0)}, \dots, \gamma^{(k)}$. Using Equation (12), we define

$$b^{(i)} = p - \frac{d\gamma^{(i)}}{dy}\Big|_{y=y^{(i)}} \tag{13}$$

and

$$\tilde{y}_j^{(i)} = \begin{bmatrix} 1 & 2y^{(i)} & \cdots & (j+1)(y^{(i)})^j \end{bmatrix} \tag{14}$$

for $i \in \{0, \dots, k\}$. We want to find the lowest order polynomial estimate of f given the data. We do this by checking if

$$\begin{bmatrix} b^{(0)} \\ \vdots \\ b^{(k)} \end{bmatrix} \in \text{range}\left(\begin{bmatrix} - & \tilde{y}_j^{(0)} & - \\ & \vdots & \\ - & \tilde{y}_j^{(k)} & - \end{bmatrix} \right) \tag{15}$$

starting with $j = 2$ and increasing it until (15) is satisfied or we reach $j = k$. Suppose that it is satisfied at $j = N$, $2 \le N \le k$. Then, we estimate $\hat{f}^{(k)}$ to be an $(N+1)$–th order polynomial. We determine α_i for $i \in \{0, \dots, N\}$ by solving

$$\begin{bmatrix} b^{(0)} \\ \vdots \\ b^{(k)} \end{bmatrix} - \begin{bmatrix} - & \tilde{y}_N^{(0)} & - \\ & \vdots & \\ - & \tilde{y}_N^{(k)} & - \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_N \end{bmatrix} = 0. \tag{16}$$

Then using $\{\alpha_i\}_{i=0}^N$, the principal solves the incentive design problem using $\hat{J}_a^{(k)}$ and $\hat{J}_p^{(k)}$. That is, the principal first solves

$$\min_{v,y} \hat{J}_p^{(k)}(v, y) = \min_{v,y} \left\{ g(y) - v + \beta \hat{f}^{(k)}(y) \right\} \tag{17}$$

subject to $v \in \mathcal{I}_v$ and $y \in \mathcal{I}_y$ to get

$$(v^{(k+1),d}, y^{(k+1),d}). \tag{18}$$

Then, the principal finds $\gamma^{(k+1)} \in \Gamma$ such that

$$\arg\max_y \hat{J}_a^{(k)}(\gamma(y), y) = y^{(k+1),d} \tag{19}$$

$$\gamma^{(k+1)}(y^{(k+1),d}) = v^{(k+1),d} \tag{20}$$

If the estimate $\hat{f}^{(k)}$ is not concave, then the solutions in the incentive design problem may only hold locally. Let us define

$$Y = \begin{bmatrix} - & \tilde{y}_N^{(0)} & - \\ & \vdots & \\ - & \tilde{y}_N^{(k)} & - \end{bmatrix}, \quad b = \begin{bmatrix} b^{(0)} \\ \vdots \\ b^{(k)} \end{bmatrix}. \tag{21}$$

from the observations $\tilde{y}_N^{(i)}$ and $b^{(i)}$ for $i \in \{0, \dots, k\}$.

Theorem 1. Suppose that f is polynomial of order $k+1$, $b \in \text{range}(Y)$, and $\text{rank}(Y) = k+1$. Then, after k iterations, the satisfaction function is known exactly and the incentive mechanism at the $(k+1)$–th step and after induces the agent to use the desired control.

Proof. Suppose that the agent's true satisfaction function is given by

$$f(y) = \sum_{i=0}^{k} \alpha_i y^{i+1}. \tag{22}$$

Let $v^d = \gamma^d(y)$ and y^d denote the desired incentive and the desired energy usage determined by maximizing

$$J_p(v, y) = g(y) - v + \beta f(y). \tag{23}$$

Since $b \in \text{range}(Y)$ and $\text{rank}(Y) = k+1$, the estimation problem defined in Equation (16) has a unique solution for $\alpha_0, \ldots, \alpha_k$ the true parameters of the agent's satisfaction function. Using f, the principal solves the incentive design problem with the true J_p and J_a to get $\gamma^{(k+1)} \equiv \gamma^d$. This incentive then induces the desired energy usage y^d. $\square$

Remark 1. The algorithm is motivated by the fact that the in the case that the agent's satisfaction function is a polynomial of order k and the principal does not know the k, but following the algorithm past even $k+1$ iterations, the principal will be playing optimally. If he had chosen the incentives $\gamma^{(i)}$ randomly, he would not know when to stop choosing a random $\gamma^{(i)}$ and thus after $k+1$ iterations would begin playing suboptimally.

We remark further that if the principal has historical data from an already existing incentive program, then the principal no longer has to go through the iterative process, i.e. if the principal has historical data on $\gamma^{(-\ell)}, \cdots, \gamma^{(1)}$ and $y^{(-\ell)}, \ldots, y^{(1)}$, then he can use this information to form a $\ell+1$ order estimate of the agent's satisfaction function.

Corollary 1. Suppose that f is polynomial up to order $k+1$ and that the principal has $k+1$ historical measurements

$$\gamma^{(-k)}, \ldots, \gamma^{(1)}, y^{(-k)}, \ldots, y^{(1)} \tag{24}$$

such that Y is full rank, then the principal can design an incentive $\gamma^{(2)}$ that induces the desired equilibrium.

We conclude this section by providing an example of the iterative process when Γ is restricted to be the space of quadratic incentives and f is a concave function.

Example 1. First, we suppose that $\gamma^{(0)}, \gamma^{(1)} \in \Gamma$ are chosen a priori and are parameterized as follows:

$$\gamma^{(i)}(y) = \xi_1^{(i)} y + \xi_2^{(i)} y^2 \tag{25}$$

for $i \in \{0, 1\}$. . Then, the procedure goes as follows. The principal issues $\gamma^{(0)}$ at time zero and observes $y^{(0)}$. Then, he issues $\gamma^{(1)}$ at time one and observes $y^{(1)}$. Using $y^{(0)}$ and $y^{(1)}$, the principal determines α_0 and α_1 in the estimate of $\hat{f}(y) = \alpha_1 y^2 + \alpha_0 y$ by computing the derivative of $\hat{J}_a^{(0)}(\gamma^{(0)}(y), y)$ and $\hat{J}_a^{(1)}(\gamma^{(1)}(y), y)$ with respect to y, evaluating at $y^{(0)}$ and $y^{(1)}$ and equating to zero, i.e.

$$-p - 2y^{(0)} + 2(\alpha_1 + \xi_2^{(0)}) y^{(0)} + \alpha_0 + \xi_1^{(0)} = 0 \tag{26}$$

$$-p - 2y^{(1)} + 2(\alpha_1 + \xi_2^{(1)}) y^{(1)} + \alpha_0 + \xi_1^{(1)} = 0 \tag{27}$$

We can solve these equations for α_1 and α_2. Then, we solve the following incentive design problem for $\gamma^{(2)}$. First, find $(v^{(2),d}, y^{(2),d})$ such that

$$\hat{J}_p^{(2)}(v, y) = -y - v + \alpha_1 y^2 + \alpha_0 y \tag{28}$$

subject to $v \in \mathcal{I}_v, y \in \mathcal{I}_y$ is maximized. Since we restrict ourselves to quadratic incentives for the sake of this example, we parameterize $\gamma^{(2)}$ as in Equation (25) with $i = 2$. Now, given the utility $\hat{J}_a^{(2)}(\gamma^{(2)}(y), y)$, we find $\xi_1^{(2)}, \xi_2^{(2)}$ such that

$$\arg\max_y \hat{J}_a^{(2)}(y; \xi_1^{(2)}, \xi_2^{(2)}) = y^{(2),d} \tag{29}$$

$$\xi_1^{(2)} y^{(2),d} + \xi_2^{(2)} (y^{(2),d})^2 = v^{(2),d} \tag{30}$$

Assuming that $y^{(2),d} \in (0, y^{\max})$, it will be an induced local maxima under the incentive $\gamma^{(2)}$. Hence, Equation (29) can be reformulated using the necessary condition

$$\nabla_y \hat{J}_a^{(2)}(y^{(2),d}; \xi_1^{(2)}, \xi_2^{(2)}) = 0. \tag{31}$$

Now, Equations (30) and (31) give us two equations in the two unknowns $\xi_1^{(2)}, \xi_2^{(2)}$ that can be solved; indeed,

$$-p + \xi_1^{(2)} + \alpha_0 + 2(\xi_2^{(2)} + \alpha_1) y^{(2),d} = 0 \tag{32}$$

$$\xi_1^{(2)} y^{(2),d} + \xi_2^{(2)} (y^{(2),d})^2 = v^{(2),d} \tag{33}$$

Solving these equations gives us the parameters for $\gamma^{(2)}$. Now, the principal can issue $\gamma^{(2)}$ to the agent and observe his reaction $y^{(2)}$. The principal can then continue in the iterative process as described above.

4. DEVICE LEVEL INCENTIVE DESIGN USING DISAGGREGATION ALGORITHM

In a manner similar to the previous section, we consider that the agent's satisfaction function is unknown. However, we now consider that the principal desires to design device level incentives. We consider two cases: the principal has an exact disaggregation algorithm and the principal has a disaggregation algorithm with bounds on the error.

4.1 Exact Disaggregation Algorithm

We first describe the process of designing device level incentives assuming the principal has a disaggregation algorithm in place which produces no error. That is, the principal observes the aggregate signal and then applies his disaggregation algorithm to get exact estimates of the device level usage y_ℓ for $\ell \in \{1, \ldots, D\}$ where D is the number of devices.

The principal has the true utility function

$$J_p(v, y) = \sum_{\ell=1}^{D} g_\ell(y_\ell) - v_\ell + \beta_\ell f_\ell(y_\ell) \tag{34}$$

and the agent has the true utility function

$$J_a(\gamma(y), y) = \sum_{\ell=1}^{D} -p y_\ell + \gamma_\ell(y_\ell) + f_\ell(y_\ell). \tag{35}$$

The principal could choose only to incentivize specific devices such as high consumption devices. This fits easily into our framework; however, for simplicity we just present the model in which incentives are designed for each device. The implicit assumption that the player utilities are separable in the devices allows us to generalize the algorithm presented in the previous section. Let us be more precise. We again assume that $\gamma_\ell^{(0)}, \gamma_\ell^{(1)}$ for $\ell \in \{1, \ldots, D\}$ are given a priori. At each iteration the principal issues an incentive for each device and observes the aggregate signal. Then he applies his disaggregation algorithm to determine the device level usage. That is to say that at the k–th iterate the principal will issue $\gamma_\ell^{(k)}$ for $\ell \in \{1, \ldots, D\}$ and observe $y^{(k)}$. Then apply a disaggregation algorithm to determine $y_\ell^{(k)}$ for $\ell \in \{1, \ldots, D\}$. The principal forms an estimate of the agent's device level satisfaction function

$$\hat{f}_\ell^{(k)}(y_\ell) = \sum_{i=0}^{j} \alpha_{i,\ell} y_\ell^{i+1}. \tag{36}$$

The principal then formulates a problem like the one in Equation (16) for each device. Let

$$y^{(i)} = (y_1^{(i)}, \ldots, y_D^{(i)}) \tag{37}$$

and

$$\gamma^{(i)} = (\gamma_1^{(i)}, \ldots, \gamma_D^{(i)}). \tag{38}$$

Then for device ℓ, given measurements $y^{(i)}$ and $\gamma^{(i)}$, we apply the condition

$$\nabla_{y_\ell} \hat{J}_a^{(i)}(\gamma(y^{(i)}), y^{(i)}) = 0 \tag{39}$$

to get

$$b_\ell^{(i)} = p - \frac{d\gamma_\ell^{(i)}}{dy_\ell}\Big|_{y_\ell = y_\ell^{(i)}}, \tag{40}$$

and

$$\tilde{y}_{\ell,j}^{(i)} = \begin{bmatrix} 1 & 2y_\ell^{(i)} & \cdots & (j+1)(y_\ell^{(i)})^j \end{bmatrix}. \tag{41}$$

We again want to find the lowest order polynomial estimate of each f_ℓ given the data. We do so by checking if

$$\begin{bmatrix} b_\ell^{(0)} \\ \vdots \\ b_\ell^{(k)} \end{bmatrix} \in \text{range}\left(\begin{bmatrix} - & \tilde{y}_{\ell,j}^{(0)} & - \\ & \vdots & \\ - & \tilde{y}_{\ell,j}^{(k)} & \end{bmatrix} \right) \tag{42}$$

starting with $j = 2$ and increasing it until (42) is satisfied or we reach $j = k$. As before, suppose that it is satisfied at $j = N_\ell$ with $2 \leq N_\ell \leq k$. Then, we estimate $\hat{f}_\ell^{(k)}$ to be an $(N_\ell + 1)$–th order polynomial. We solve

$$\tilde{b}_\ell - Y_\ell \tilde{\alpha}_\ell = \begin{bmatrix} b_\ell^{(0)} \\ \vdots \\ b_\ell^{(k)} \end{bmatrix} - \begin{bmatrix} - & \tilde{y}_{\ell,N_\ell}^{(0)} & - \\ & \vdots & \\ - & \tilde{y}_{\ell,N_\ell}^{(k)} & - \end{bmatrix} \begin{bmatrix} \alpha_{0,\ell} \\ \vdots \\ \alpha_{j,\ell} \end{bmatrix} = 0 \tag{43}$$

for $\alpha_{0,\ell}, \ldots, \alpha_{j,\ell}$. We repeat this process for each device $\ell \in \{0, \ldots, D\}$. Using the now estimated $\hat{f}_\ell^{(k)}$ for each device, the principal solves the incentive design problem for $\gamma_\ell^{(k+1)}$, $\ell \in \{1, \ldots, D\}$. He does so by first solving

$$(v^{(k+1),d}, y^{(k+1),d}) = \arg\max_{v,y} \hat{J}_p^{(k)}(v, y) \tag{44}$$

where $v = (v_1, \ldots, v_D)$, $y = (y_1, \ldots, y_D)$, $v_\ell \in \mathcal{I}_{v_\ell}$, $y_\ell \in \mathcal{I}_{y_\ell}$ and

$$\hat{J}_p^{(k)}(v, y) = \sum_{\ell=1}^{D} g_\ell(y_\ell) - v_\ell + \beta_\ell \hat{f}_\ell^{(k)}(y_\ell). \tag{45}$$

Then, the principal finds $\gamma_\ell^{(k+1)} \in \Gamma$ such that

$$\arg\max_y \hat{J}_a^{(k)}(\gamma(y), y) = y^{(k+1),d} \tag{46}$$

$$\gamma^{(k+1)}(y^{(k+1),d}) = v^{(k+1),d} \tag{47}$$

where $\gamma^{(k+1)} = (\gamma_1^{(k+1)}, \ldots, \gamma_D^{(k+1)})$ and

$$\hat{J}_a^{(k)}(\gamma(y), y) = \sum_{\ell=1}^{D} -py_\ell + \gamma_\ell^{(k+1)}(y_\ell) + \hat{f}_\ell^{(k)}(y_\ell). \tag{48}$$

This process is repeated as in the previous section.

Theorem 2. Suppose that each f_ℓ for $\ell \in \{1, \ldots, D\}$ is polynomial up to order $k_\ell + 1$, $\tilde{b}_\ell \in \text{range}(Y_\ell)$, and $\text{rank}(Y_\ell) = k_\ell + 1$. Then, after

$$k^* = \max_{\ell \in \{1, \ldots, D\}} k_\ell \tag{49}$$

iterations, the satisfaction function is known exactly and the incentive mechanism at the $(k^* + 1)$–th step and after induces the agent to use the desired control.

The proof of the theorem is similar to that of Theorem 1.

4.2 Disaggregation Algorithm with Some Error

Now, we consider that the principal has some error in his estimate of the device level usage due to inaccuracies in the disaggregation algorithm, i.e. the principal determines $\hat{y}_\ell$ such that $\|y_\ell - \hat{y}_\ell\| \leq \varepsilon$ where $\varepsilon > 0$ is the resulting error from the estimation in the disaggregation algorithm. Bounds on ε can be determined by examining the fundamental limits of non–intrusive load monitoring algorithms (Dong et al., 2013b).

The incentive design process follows the same scheme as provided in the previous subsection with the exception that in the disaggregation step the estimate of y_ℓ is not exact. We again assume that $\gamma_\ell^{(0)}, \gamma_\ell^{(1)}$ for $\ell \in \{1, \ldots, D\}$ are given a priori. Following the same procedure as before, at the k–th iterate the principal will issue $\gamma_\ell^{(k)}$ for $\ell \in \{1, \ldots, D\}$ and observe $y^{(k)}$. Then apply a disaggregation algorithm to determine $\hat{y}_\ell^{(0)}$ where

$$\|y_\ell - \hat{y}_\ell\| \leq \varepsilon \tag{50}$$

for $\ell \in \{1, \ldots, D\}$. The incentive design problem follows the same steps as provided in Equation (36)–(48) with the exception that the y_ℓ's are replaced with the estimated $\hat{y}_\ell$'s and we tolerate an error in solving (43) for the minimal polynomial estimate of f_ℓ.

5. NUMERICAL EXAMPLES

We simulate two examples of designing incentives while estimating the agent's satisfaction function. The first example consists of an agent with a concave satisfaction function that is not quadratic. We design incentives based on the aggregate energy signal. The second example consists of an agent having a quadratic satisfaction function and the principal designs incentives using the disaggregated energy signal where the disaggregation algorithm has some error. In both examples we use a price of $p = 1$.

5.1 Aggregate Signal and Log Satisfaction Function

We simulate a system in which the agent has the true utility given by

$$J_a(\gamma(y), y) = -py + \gamma(y) + f(y) \tag{51}$$

where the satisfaction function is

$$f(y) = 10 \log(y + 1). \tag{52}$$

The principal's cost function is

$$J_p(v, y) = -y - v + \beta f(y) \tag{53}$$

where the benevolence factor is $\beta = 0.75$. We let $y^{\max} = v^{\max} = 100$. We choose two concave incentive function $\gamma^{(0)}(y)$ and $\gamma^{(1)}(y)$ defined as follows:

$$\gamma^{(0)}(y) = -y^2 + 10y, \;\; \gamma^{(1)}(y) = -y^2 + 15y. \tag{54}$$

We use the algorithm presented in Section 3 to design incentives while estimating α_0 and α_1. We simulate the principal issuing $\gamma^{(0)}$ and then $\gamma^{(1)}$ where the agent chooses his

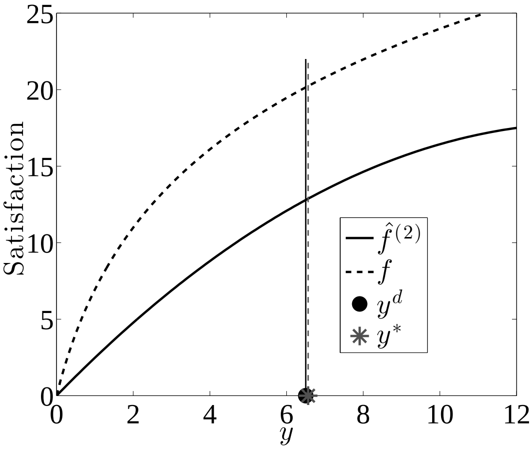

Fig. 2. Estimated satisfaction function $\hat{f}^{(2)}$ and true satisfaction function f. The true response $y^* = 6.56$ and the desired response $y^d = 6.5$. Notice that the slope of the estimated satisfaction function and the slope of the true satisfaction function are roughly equal at y^d and y^*.

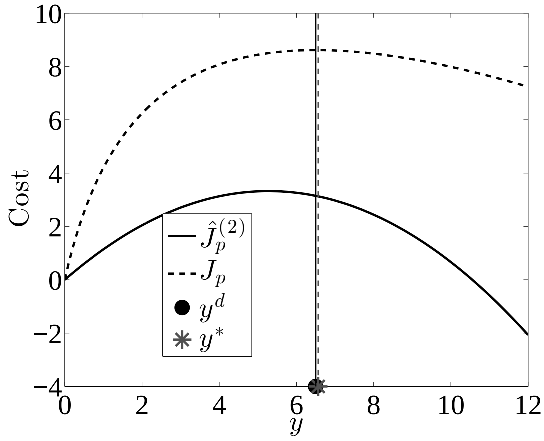

Fig. 3. Estimated cost function $\hat{J}_p^{(2)}(\gamma^{(2)}(y), y)$ and true cost function $J_p(v^d, y)$ along $v = v^d$ for simulation with log–satisfaction function. Note that the *shape* of $J_p(v^d, y)$ and the *shape* of $\hat{J}_p^{(2)}(\gamma^{(2)}(y), y)$ are the similar.

optimal response to each of the incentives. The responses are $y^{(0)} = 5.29$ and $y^{(1)} = 7.58$. After two iterations, we get a reasonable approximation of the true f and a quadratic incentive $\gamma^{(2)}$;

$$\hat{f}^{(2)}(y) = 2.57y - 0.093y^2, \ \gamma^{(2)}(y) = 0.33y - 0.05y^2. \quad (55)$$

The optimal power usage under the incentive $\gamma^{(2)}$ is $y^* = 6.56$ and the desired power usage is $y^d = 6.5$. It is clear that the principal could do better if he new the true satisfaction function. Figure 2 shows $\hat{f}^{(2)}(y)$ and $f(y)$. It is important to notice that y^* is nearly equal to y^d and at these two points the slope of $\hat{f}^{(2)}$ is approximately equal to that of the true f. This indicates that $\hat{f}^{(2)}$ is a good estimate of f. Figure 3 shows the true utility function of the principal $J_p(v^d, y)$ with $v = v^d$ fixed and the estimated utility $\hat{J}_p^{(2)}(\gamma^{(2)}(y), y)$. y^d is the point at which $J_p(v^d, y)$ is maximized and it is approximately the point where $\hat{J}_p^{(2)}(\gamma^{(2)}(y), y)$ is maximized. It is important to note the *shape* of J_p and $\hat{J}_p^{(2)}$. The offset is not important because we are not estimating a constant term in $\hat{f}$ since it does not affect the optimal response, i.e. if you shift $\hat{J}_p^{(2)}$ by a constant term, y^* is still the optimal response. In Figure 3, J_p and $\hat{J}_p^{(2)}$ have a similar shape.

5.2 Disaggregated Signal

We simulate a system in which the agent the true utility

$$J_a(\gamma(y), y) = \sum_{i=1}^{10} -py_i + \gamma_i(y_i) + f_i(y_i) \quad (56)$$

where the satisfaction functions $f_i(y_i)$ are exactly quadratic for each device $i \in \{1, \ldots, 10\}$;

$$f_i(y_i) = \alpha_{1,i}y_i^2 + \alpha_{0,i}y_i \quad (57)$$

The principal's utility is given by

$$J_p(v, y) = \sum_{i=1}^{10} -y_i - v_i + \beta_i f_i(y_i) \quad (58)$$

where the benevolence factor is $\beta_i = 1$ for each i. The principal must disaggregate the aggregated energy signal y giving rise to estimates $\hat{y}_i$. If $\hat{y}_i = y_i$, i.e. there is no error in the disaggregation algorithm, then after two iterations the principal would know the satisfaction function of each device exactly. Let explore the case when the disaggregation algorithm has ε–error. In our examples we randomly generate noise within a given ε bound and add that to the true y_i's to simulate the error in the disaggregation step resulting from the disaggregation algorithm.

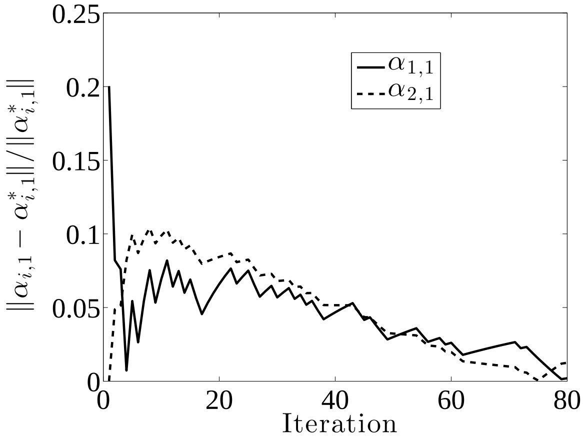

Fig. 4. Relative error in estimate of $\alpha_{i,1}$'s for device 1 with disaggregation error bound $\varepsilon = 0.15$. α_i^* is the true value. The relative error eventually decreases below the noise bound $\varepsilon = 0.15$.

Figure 4 shows the relative error on the estimates of $\alpha_{i,1}$ for $i \in \{1, 2\}$ as a function of the iteration. The relative error for other devices are similar. We used the error bound $\varepsilon = 0.15$ for the error from disaggregation. The relative error decreases as the number of iterations increase. It eventually ends up below the noise bound ε and remains there.

As we iterate the noise introduced via disaggregation has minimal impact on the estimate of $\alpha_{i,\ell}$ for $i = \{1, 2\}$ and $\ell \in \{1, \ldots, D\}$. We note that the designed incentive for this problem converges to zero as we increase the iterations and the impact of the noise is minimized. It becomes zero since the benevolence factor is $\beta_\ell = 1$ and the price $p = 1$; hence, the agent and the principal have the same utility functions after the principal learns the agent's satisfaction function. As we increase the noise threshold ε, the estimation of $\alpha_{i,\ell}$ degrades.

In the last simulation, we decrease the benevolence factor to $\beta = 0.75$ and comment on the resulting incentives. In

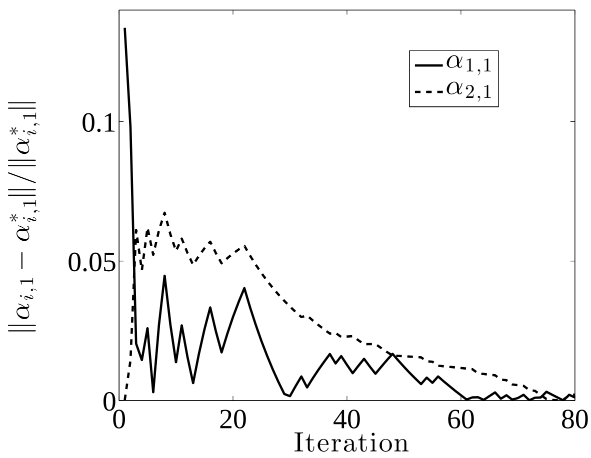

Fig. 5. Relative error in estimate of $\alpha_{i,1}$'s for device 1 with disaggregation error bound $\varepsilon = 0.1$. α_i^* is the true value. The relative error eventually decreases below the noise bound $\varepsilon = 0.1$;

Figure 5 we show again the relative error in the estimates of $\alpha_{i,1}$ for $i \in \{1, 2\}$. The relative error of each of the other devices is similar. In this case the incentive for device $\ell = 1$ converge to

$$\gamma_1^*(y) = -0.39y^2 + 0.33y. \quad (59)$$

The other devices have similar incentives. The reason that there is a non–zero incentive is due to the fact that the principal is not completely benevolent; he does not care as much about the satisfaction of the agent as he does the other terms in his cost function. However, as we iterate and the principal learns the agent's cost function, the principal is able to use the incentives to force the desired action $y^* = y^d$ where y^* is the agent's true response.

6. DISCUSSION AND FUTURE WORK

We modeled the utility company–consumer interaction as a principal–agent problem in which the utility company is the principal and the consumer is the agent. We defined a process by which the utility company can jointly estimate the agent's utility function and design incentives for behavior modification. We solved this problem for both the case where the principal designs incentives using the aggregate consumption and the case when the principal has a disaggregation algortihm in place and designs device level incentives. Whether the principal is interested in inducing energy efficient behavior or creating an incentive compatible demand response program, the procedure we present applies.

We are studying fundamental limits of *non–intrusive load monitoring* in order to determine precise bounds on the payoff to the utility company when a disaggregation algorithm is in place and incentives are being designed. In addition, we seek to understand how these fundamental limits impact the quality of the incentive design problem as a whole. Our algorithm provides a means to estimate the satisfaction function while designing incentives and is motivated by the fact that the principal is using his belief garnered from the information availble to him in order to simulatenously do control and estimation. There are some limitations to the algorithm we propose. The necessary condition, namely the zero gradient of the agent's cost function evaluated at the agent's response, we use to formulate the estimation problem for the satisfaction function holds only on the interior of the domain. We are studying how to handle the boundary conditions on constraints within our proposed algorithm.

The electrical grid is a cyber-physical system with human actors influencing the trajectory of the system. This is sometimes referred to as a human cyber-physical system (h-CPS). Casting behavior modification in the electrical grid as a control and estimation problem allows us to explicitly model the human components and the cyber-physical components in one framework. Inherent to the study of h-CPS's are privacy and security considerations. We remark that consumers may consider their satisfaction function to be private information. We are currently exploring the design of privacy–aware mechanisms for ε–incentive compatible problems (Nissim et al., 2012). Further, we are currently studying security consideration by modeling *adversarial consumers* who may wish to spoof their aggregated signal or *adversarial external agents* who wish to comprise the system resulting in catastrophic failure of demand response programs dependent on the disaggregation based incentives.

REFERENCES

Armel, K.C., Gupta, A., Shrimali, G., and Albert, A. (2013). Is disaggregation the holy grail of energy efficiency? The case of electricity. *Energy Policy*, 52, 213–234.

Basar, T., Olsder, G.J., Clsder, G., Basar, T., Baser, T., and Olsder, G.J. (1995). *Dynamic noncooperative game theory*, volume 200. SIAM.

Bertsekas, D.P. (1999). *Nonlinear programming.* Athena Scientific.

Creyts, J., Derkach, A., Nyquist, S., Ostrowski, K., and Stephenson, J. (2007). Reducing U.S. greenhouse gas emissions: How much at what cost? Technical report, U.S. Greenhouse Gas Abatement Mapping Initiative.

Dong, R., Ratliff, L., Ohlsson, H., and Sastry, S.S. (2013a). Energy disaggregation via adaptive filter. In *Proceedings of the 50th Allerton Conference on Communication, Control, and Computing*.

Dong, R., Ratliff, L., Ohlsson, H., and Sastry, S.S. (2013b). Fundamental limits of non–intrusive load monitoring. *arXiv:1310.7850v1*.

Eom, J. (2008). Shareholder incentives for utility-delivered energy efficiency programs in California. In *Proceedings of the 28th USAEE/IAEE North American Conference*.

Froehlich, J., Larson, E., Gupta, S., Cohn, G., Reynolds, M., and Patel, S. (2011). Disaggregated end-use energy sensing for the smart grid. *IEEE Pervasive Computing*, 10(1), 28–39.

Gardner, G.T. and Stern, P.C. (2008). The short list: The most effective actions U.S. households can take to curb climate change. In *Environment: Science and Policy for Sustainable Development*.

Gupta, S., Reynolds, M.S., and Patel, S.N. (2010). Electrisense: single-point sensing using EMI for electrical event detection and classification in the home. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, Ubicomp '10, 139–148. ACM, New York, NY, USA.

Johnson, M.J. and Willsky, A.S. (2012). Bayesian nonparametric hidden semi-Markov models. *arXiv:1203.1365*.

Laffont, J.J. and Martimort, D. (2009). *The theory of incentives: the principal-agent model*. Princeton University Press.

Laitner, J.A., Ehrhardt-Martinez, K., and McKinney, V. (2009). Examining the scale of the behaviour energy efficiency continuum. In *European Council for an Energy Efficient Economy*.

Leeb, S., Shaw, S., and Kirtley, J.L., J. (1995). Transient event detection in spectral envelope estimates for nonintrusive load monitoring. *IEEE Transactions on Power Delivery*, 10(3), 1200–1210.

Mathieu, J.L., Haring, T., and Andersson, G. (2012). Harnessing residential loads for demand response: Engineering and economic considerations. In *Interdisciplinary Workshop on Smart Grid Design and Implementation*.

Nissim, K., Orlandi, C., and Smorodinsky, R. (2012). Privacy–aware mechanism design. *arXiv:1111.3350v2*.

Perez-Lombard, L., Ortiz, J., and Pout, C. (2008). A review on buildings energy consumption information. *Energy and buildings*, 40, 394–398.