52nd IEEE Conference on Decision and Control
December 10-13, 2013. Florence, Italy

On the optimality of de-synchronized demand response with stochastic renewables and inertial thermal loads

Gaurav Sharma[1], Le Xie[1] and P. R. Kumar[1]

Abstract— **We consider the problem of optimally supporting a collection of thermostatically controlled inertial loads in a smart microgrid environment. One key consideration is to reduce the need for costly nonrenewable reserves, while also meeting the consumer requirements specific to inertial thermal loads at the same time. Several models studied for optimal demand response leads to a response that is synchronized over the collection of loads, which is undesirable since synchronized response entails high demand fluctuations. We propose a model with an additional feature of stochastic variations in end user requirements. We show that employing a cost function that additionally penalizes the error in adhering to end user requirements leads to an optimal demand response that is de-synchronized over the collection of homogeneous loads. That is, identical thermal loads are intentionally staggered so as to hedge against future uncertainty of consumer preferences. We show that such de-synchronization is related to the concavity of the cost-to-go function in the Hamilton-Jacobi-Bellman equation. We provide a simple heuristic to approximate the optimal policy. We also illustrate the results on a simple numerical example.**

I. INTRODUCTION AND MOTIVATION

There has been a steady growth in the usage of renewable green energy sources over the last few years [13],[15]. The growth of these alternative generation sources is not only fueled by their reduced carbon footprint, but also by strong policy incentives to ultimately provide low cost power from renewables. It is estimated that short-run variable cost of wind electricity production is 5-8 cents per kWh range, which is more economic than fossil fuel generation [14].

However, generation from these sources is inherently variable, and therefore not reliable for dispatching. To compensate for this randomness in generation, demand response in the form of load control is preferable over generation based compensation because of two reasons, 1) the inherent flexibility in loads could allow flexible ramping capabilities from demand side rather than costly ramping of generators, and 2) in the deregulated markets, generators need to be committed much ahead of time which is both costly as well as requires prediction of future demand, and is therefore subject to error in prediction. To implement such nimble demand response, new smart grid technologies are fostering an infrastructure of smart devices which can be controlled over a communication infrastructure.

In this paper, our main focus is on thermal inertial loads such as air-conditioners, with controllable thermostats, which can be observed and controlled over the supporting communication network.

We consider a load serving entity (LSE) which is connected to both a renewable power driven micro-grid as well as the main power grid. The LSE can both observe and control a collection of thermostatically controllable loads (TCLs), and draw the appropriate wind and grid power combination to satisfy its objective. We analyze several optimization criteria that attempt to take into account both end user performance (e.g., comfort) as well as the need to reduce variations in the grid power drawn. Cost criteria are however ultimately only a tool for obtaining suitable control policies via stochastic optimal control theory. In fact, it turns out that not all cost criteria that appear reasonable upon first glance are in actuality sound. The difference is in their robustness to violations in the model assumptions. The key issue we seek to understand is the robustness of the resultant optimal control law in terms of the variations in the power drawn from the grid. We show that this involves several subtleties, and propose a suitable cost criterion which incorporates additional modeling features to ensure robustness.

Our main results speak to the nature of the optimal response for various choices of cost criteria. We show that under the homogeneous loads case, the optimal response under several reasonable looking cost criteria that explicitly penalize the variations in grid power drawn is to synchronize the temperatures of all the loads. At a mathematical level, this synchronization results from the convexity of the cost-to-go function satisfying the Hamilton-Jacobi-Bellman (HJB) equation. However, while this synchronized response is optimal in terms of balancing end-user comfort versus variations in grid power drawn, it has a serious flaw. If the end-users were to synchronously change their temperature set-point, as for example may happen on Super Bowl Sunday at gametime when everyone at home gathers in the same room, the resulting response exhibits undesirable oscillations in grid power drawn. To address this consequence of an otherwise reasonable cost function, we incorporate an additional model feature – stochastic variation of the common set-point across the collection of TCLs. When such synchronized randomness in user comfort is introduced in the model we show that the resulting optimal control law intentionally separates the temperatures of the inertial loads by giving preference to certain loads over others. That is, a deliberately de-synchronized solution is optimal. Technically, this symmetry breaking is associated with the local concavity of the cost-to-go function satisfying the HJB equation.

[1]Department of Electrical and Computer Engineering, Texas A & M University. Email:{gash,le.xie,prk}@tamu.edu.

[2]This material is based upon work partially supported by NSF under Contract Nos. CPS-1239116, CPS-1232601, CPS-1232602 and ECCS-1150944.

II. RELATED WORK

There has been much research into modeling, designing and analyzing demand responses for thermostatically controlled loads. Lu et. al. [7] have used a queueing model to analyze the diversity in the states of thermal loads, which can be lost by the undesired price-based demand response of aggregated TCLs. In [10], Lu and Katipamula have discussed control strategies for adjusting thermostat set-points of TCLs for different energy price-periods to reduce peak power consumption and energy cost. Other works have focused on improving the stability of the power system by a TCL-based ancillary service. Xu et. al. [4] have designed a control logic and a model for frequency activated thermal loads which can vary their set-points to provide additional reserves for system stability. Short et. al. [5] discuss a model in which TCLs monitor the system's frequency and toggle between *on* and *off* depending on a degree of compromise between the need of the appliance and that of the grid.

Callaway [3] has proposed a method for controlling loads via programmable thermostats to produce short time scale responses to deliver load following and regulation in the grid. Kundu et.al. [2] have modeled aggregate power response of set-point variation by a homogeneous population of TCLs. To prevent the resulting oscillations they subsequently propose a LQR regulator on the linearized model. Kundu and Sinitsyn [1] have developed a protocol to safely shift the set-points to mitigate the fluctuations in power drawn. Our analysis differs from preceding works mainly in two ways. First, we consider inertial loads as buffers to absorb the fluctuations of the renewable generators, and second, rather than proposing an ad-hoc approach to obtain the staggering of the inertial loads, we attempt to answer a more fundamental yet subtle question: Why is the staggering desirable in optimization based demand response?

III. MODELS, OPTIMIZATION CRITERIA & RESULTING DEMAND RESPONSES

We address the problem of centrally controlling the thermostatically controlled loads (TCLs) in the presence of stochastic renewable sources. For specificity, we assume that the renewable energy source is wind. We will refer to the central controlling agent as a load serving entity (LSE). The LSE serves two functions. First, it estimates and allocates the available renewable power $W(t)$ along with 'grid' power (by which we mean the traditional non-renewable sources) to individual TCLs. Second, the LSE has a separate communication channel to monitor each TCL's state and send the control actions to all the loads. The LSE can utilize the knowledge of the states of the TCLs as well as the renewable generation process to achieve predefined objectives. From the power distribution grid side the LSE acts as a cumulative load which co-ordinates TCLs efficiently.

The end-users are assumed to have some flexibility, specified as a continuous range of acceptable temperatures for TCLs. From the grid's perspective end-users can be incentivized to provide more margin to compensate for stochasticity in renewable power generation, and in effect reduce grid power consumption.

For specificity, we consider the problem of cooling. In our models we make several simplifications to focus on the central problem of interest, which is to understand how the de-synchronization emerges naturally from an optimization objective imposed on a model. First , while in practice, residential thermal loads have a constant cooling rate and consume fixed power, we relax this and assume that loads can be run at fractional capacity. This fractional cooling rate can be approximated by fast switching (chattering) between two cooling rates. Second, we simplify analysis by assuming constant ambient heating, resulting in the fact that the rate of change in temperature is assumed to be an affine function of power supplied, i.e., temperature dynamics of $\dot{x} = S - P_{Load}$, which we shall refer to as "constant dynamics." However the results can be easily extended to any linear dynamics, e.g., to $\dot{x} = Ax + S - P_{Load}$.

In the sequel we will consider several end-user criteria. One possibility is to measure end-user discomfort by the square of the amount by which the temperature of the TCL exceeds a user specified maximum temperature Θ_M. This corresponds in an optimization framework to a cost function $[(x - \Theta_M)^+]^2$, where x is the temperature of a TCL and $z^+ := \text{Max}(z, 0)$. Another possibility is to impose a hard upper bound constraint Θ_M. Then, in order to keep the temperature below Θ_M, the TCL may have to draw grid power if wind power is not available. In the simple model $\dot{x} = S - P_{load}$, if x hits Θ_M, and the wind is not blowing, the TCL will have to draw power of at least S. The comfort setting Θ_M could also be stochastic. In that case, we require that if $x > \Theta_M$, then the TCL must draw maximum grid power M, where $M > S$. We may also have a minimum temperature constraint Θ_m. Then, when x hits Θ_m, the total of grid and wind power chosen must not be greater than S.

Concerning reducing variations in the power P^g drawn from the grid, we will penalize variations in it by employing a quadratic running cost $\int_0^T (P^g(t))^2 dt$, where $P^g(t) := \sum_i P_i^g(t)$ is the sum of the powers drawn by the individual TCLs. In place of the quadratic penalty, any other strictly increasing, strictly convex penalty in the total grid power drawn would also penalize variations.

Turning to the model of the wind process, there are several choices in the literature [11],[12]. For simplicity, we will focus on a simplified wind process modeled as a two state continuous time Markov chain. States *on* and *off* have mean holding times $1/q_1$ and $1/q_0$ respectively. We will refer to this process as a $\mathcal{M}(q_0, q_1)$ process. In the *on* state the wind can provide W watts, while there is zero wind power in the *off* state.

By a *synchronizing* policy we will mean a policy which reduces the difference in TCL temperature states by allocating more cooling power to a higher temperature TCL. On the contrary a *de-synchronizing* policy may occasionally increase the temperature difference by providing power to a lower temperature TCL and thus cooling it, while letting the higher temperature one increase. We note that a de-synchronizing policy may be a state dependent policy that sometimes drives temperatures towards one another, and sometimes apart, i.e., it need not always strive to separate temperatures.

Now we describe several alternative models and optimization criteria, and the nature of demand response that results from the combination of both. The proofs of all the results are deferred to Section IV.

A. Contract violation probability (CVP) model

One model of a contract with end users is to maintain their temperatures in a specified comfort range. This results in an objective of minimizing the probability of contract violation. Consider N TCLs with temperatures denoted $(x_1, x_2, ..., x_N)$, following the temperature dynamics $\dot{\vec{x}} = \vec{f}(\vec{P})$ as a function of supplied power $\vec{P}$. Suppose the contractual ranges are $\{[\Theta_m^i, \Theta_M^i]\}_{i=1}^N$. Then we obtain the following optimization problem to minimize the contract violation probability:

$$\textbf{Minimize} \sum_{i=1}^{N} \text{Prob } (\{x_i \notin [\Theta_m^i, \Theta_M^i]\})$$

$$\textbf{Subject to } \frac{d\vec{x}}{dt} = \vec{f}(\vec{P}(t)) \tag{1}$$

$$\sum_{i=1}^{N} P_i(t) \leq W(t) \tag{2}$$

$$P_i(t) \geq 0 \text{ for } i = 1, 2, ..N \tag{3}$$

$$W(t) \backsim \mathcal{M}(q_0, q_1). \tag{4}$$

For this problem, we will show in the sequel that since all loads outside the comfort zone contributes the same amount to cost, the loads which are nearer to the comfort zone are preferred by the optimization policy because they can be brought within the comfort range more quickly. Thus the resulting scheme is not fair across loads.

B. Variance minimization model

To correct the above bias we propose next a model in which the penalty function is the square of the deviation above Θ_M. The optimization problem is:

$$\textbf{Minimize} \int_0^T \Big(\sum_{i=1}^{N} \mathbb{E}[(x_i(t) - \Theta_m^i)^+]^2) | \vec{x}(0)] \Big)$$

$$\textbf{Subject to } \text{ Constraints (1-4)}.$$

Due to the convex quadratic penalty, keeping the TCLs' temperatures apart from one another costs more than keeping their temperatures the same. The result is that it is in fact optimal to bring all TCLs to the same temperature and maintain them so.

C. Hard temperature threshold model

Notice that the preceding model does not guarantee any upper bound on the maximum temperature level. Such a hard temperature constraint cannot be met without a reliable grid power source. Therefore, we introduce a hard temperature upper bound in addition to the soft temperature comfort goal mentioned earlier. We also allow for but penalize the grid power required to operate within the hard bound. This yields:

$$\textbf{Min. } \mathbb{E}\Big[\int_0^T \Big(\sum_{i=1}^{N} [(x_i(t) - \Theta_{soft}^i)^+]^2 + \gamma \Big(\sum_{i=1}^{N} P_i^g(t) \Big)^2 | \vec{x}(0) \Big) \Big]$$

$$\textbf{Subject to } \frac{d\vec{x}}{dt} = \vec{f}(\vec{P}^w(t) + \vec{P}^g(t)) \tag{5}$$

$$x_i(t) \in [\Theta_m^i, \Theta_M^i] \text{ for } i = 1, 2, ..., N \tag{6}$$

$$\sum_{i=1}^{N} P_i^w(t) \leq W(t) \tag{7}$$

$$P_i^w(t) \geq 0, P_i^g(t) \geq 0 \text{ for } i = 1, 2, ...N \tag{8}$$

and (4).

In this model, at first sight one may speculate that keeping the TCLs apart hedges against their hitting the maximum temperature at the same time, thereby reducing the term in the cost function that is quadratic in the total power. However, it turns out that driving the TCLs to the same temperature and then maintaining them at equal temperature continues to be optimal, as we show in Section V. So, we conclude that this model does not result in the desired behavior of desynchronization.

D. Stochastic threshold variation (STV) model

We now introduce the additional feature that there could be environmental, social or other extraneous events due to which the end user may change the set-point in a coordinated fashion. To capture this effect we will assume that there are two levels Θ_M^1 and Θ_M^2, where $\Theta_M^1 < \Theta_M^2$. All the TCLs switch between these levels at the same time instants according to a process $Th(t) \backsim \mathcal{M}(r_{high}, r_{low})$, i.e. as a Markov process with mean holding times $1/r_{high}$ and $1/r_{low}$ in the two states Θ_M^2 and Θ_M^1. Due to a sudden reduction in the set-point, a TCL that was previously within the desired temperature range $[\Theta_m, \Theta_M^2]$ may suddenly be at a higher temperature than Θ_M^1. When a TCL thereby violates the threshold constraint we will require that it be provided grid power at the maximum possible level M that the TCL can sustain, to cool it quickly:

$$\textbf{Min. } \int_0^T \mathbb{E}\Big[\Big(\sum_{i=1}^{N} P_i^g(t) \Big)^2 | \vec{x}(0) \Big]$$

$$\textbf{s.t. } \text{ (5,7,8) and}$$

$$x_i(t) \in [\Theta_m, \Theta_M^2] \text{ for } i = 1, 2, ..., N$$

$$P_i^g = M \ \ \{ \text{ if } x_i(t) > Th(t) \}$$

$$W(t) \backsim \mathcal{M}(q_0, q_1), Th(t) \backsim \mathcal{M}(r_{high}, r_{low}).$$

We will show that with this model the optimal power allocation necessarily results in de-synchronization of the TCL temperature states. The optimal policy does not maintain all the TCLs at the same state. When there are TCLs above a certain level it is optimal to keep their temperatures different, to hedge against the future eventuality that the thermostats are switched down to Θ_M^1.

IV. RESULTS

For simplicity, we will consider a homogeneous population of TCLs, with identical temperature ranges, dynamics, heating, and user preference variations. We illustrate all results, except the numerical computation for the $N = 2$ case.

Theorem 1: The optimal response in the CVP model with $\mathcal{M}(q_0, q_1)$ wind process and linear dynamics is to provide all the wind power to the coolest TCL that is outside the temperature range. Therefore the optimal policy is *de-synchronizing*.

Theorem 2: Under the variance minimization model and $\mathcal{M}(q_0, q_1)$ wind process, the optimal control policy is of *synchronizing* nature, for any linear dynamics.

Theorem 3: Under the hard temperature threshold model with $\mathcal{M}(q_0, q_1)$ wind process and constant dynamics the optimal policy is of synchronizing nature.

Last, we illustrate the *de-synchronizing* nature of the optimal policy for the *stochastic threshold variation* model, by numerically computing its solution.

V. PROOFS

Proof of Theorem 1: For N=2 and a wind power realization $W(t, \omega)$, and, for initial state $\vec{x}(0) = (x_1(0), x_2(0))$ with $\Theta_M < x_1(0) < x_2(0)$, let the optimal cost be $C^*_\omega(\vec{x}(0))$, and the resulting state $\overset{\star}{x}(t, \omega)$. Now consider a policy $\hat{\Pi}$, with resulting cost $\hat{C}_\omega(\vec{x}(0))$, which gives all the power to TCL1 while $x_1 > \Theta_M$, and subsequently to TCL2 while maintaining TCL1 at Θ_M, and denote the resulting state by $\hat{x}(t, \omega)$. Due to linear dynamics we have $\overset{\star}{x}_1(t, \omega) + \overset{\star}{x}_2(t, \omega) = \hat{x}_1(t, \omega) + \hat{x}_2(t, \omega)$, therefore the event $\{\hat{x}_1, \hat{x}_2 > \Theta_M\} \subset \{\overset{\star}{x}_1, \overset{\star}{x}_2 > \Theta_M\}$. Also the probability of the event $\{\hat{x}_i > \Theta_M, \hat{x}_{!i} \leq \Theta_M\}$, where $!i$ denotes the TCL other than TCL i, is minimized by $\hat{\Pi}$ we have,

$$C^*_\omega(\vec{x}(0)) = \int_{\{\overset{\star}{x}_i > \Theta_M, \overset{\star}{x}_{!i} \leq \Theta_M\}} 1dt + \int_{\{\overset{\star}{x}_1, \overset{\star}{x}_2 > \Theta_M\}} 2dt$$
$$\geq \int_{\{\hat{x}_i > \Theta_M, \hat{x}_{!i} \leq \Theta_M\}} 1dt + \int_{\{\hat{x}_1, \hat{x}_2 > \Theta_M\}} 2dt = \hat{C}_\omega(\vec{x}(0)).$$

Taking the expectation yields the desired result. ■

A. HJB equation for variance minimization model

Let the optimal cost-to-go from state (x_1, x_2) and wind condition i (i=1 in *on* state and i=0 on *off* state) be $V_i^*(\vec{x}, t)$. The state-space is $\mathcal{S} = [0, \infty) \times [0, \infty) \times \{0, 1\}$. We denote by $\mathcal{A} \subset \mathbb{R}^2_+$ be the set of state-dependent *admissible* actions

$$\mathcal{A}(\vec{x}, i) = \{(P_1, P_2) : P_1 + P_2 \leq W \text{ if (wind } i = 1),$$
$$P_1 = P_2 = 0 \text{ if (wind } i = 0),$$
$$\text{if}(x_j(t) = 0) \text{ then } f(P_j) \geq 0\}.$$

The Hamilton-Jacobi-Bellman equation is:

$$\inf_{(P_1,P_2) \in \mathcal{A}(\vec{x}, i)} \{\mathbf{1}^T[(\vec{x} - \Theta_M)^+]^2 + \nabla^T V_i^*(\vec{x}, t)\vec{f}(\vec{P}) - q_i V_i^*(\vec{x}, t)$$
$$+ q_i V_{!i}^*(\vec{x}, t)\} + \frac{\partial V_i^*}{\partial t}(\vec{x}, t) = 0 \quad \text{For } i = 0, 1. \quad (9)$$

Equation (9) involves the following minimization problem:

$$\inf_{(P_1,P_2) \in \mathcal{A}(\vec{x}, i)} \left\{ \frac{\partial V_i^*}{\partial x_1}(\vec{x}, t) f(P_1) + \frac{\partial V_i^*}{\partial x_2}(\vec{x}, t) f(P_2) \right\}. \quad (10)$$

Since $f(P)$ is linear, the infimum above is achieved for either $P_1 = W$ or $P_2 = W$, depending on whether $\frac{\partial V_i^*}{\partial x_1} > \frac{\partial V_i^*}{\partial x_2}$ or $\frac{\partial V_i^*}{\partial x_1} < \frac{\partial V_i^*}{\partial x_2}$.

Lemma 1: The optimal cost-to-go function $V^*(\vec{x}, t)$ is component-wise increasing in $\vec{x}$.

Proof: For a realization of wind power $W(t, \omega)$, let the optimal state trajectory from initial state $(x_1(0), x_2(0))$ be $(x_1(t, \omega), x_2(t, \omega))$. From initial state $y_1(0) < x_1(0)$, we provide the same power as before, except at times when temperature goes below Θ_m to obtain trajectory $(\hat{y}_1(t, \omega), x_2(t, \omega))$. Where $\hat{y}_1(t, \omega) = max\{\Theta_m, y_1(0) - x_1(0) + x_1(t, \omega)\}$. We have,

$$V^*_\omega(\vec{x}(0)) \geq \int_0^T [(\hat{y}_1(t, \omega) - \Theta_M)^+]^2 + [(x_2(t, \omega) - \Theta_M)^+]^2 dt$$
$$= \hat{V}_\omega(x_1(0), x_2(0)) \geq V^*_\omega(x_1(0), x_2(0)).$$

Taking expectation yields the desired result. ■

Lemma 2: If $f(P)$ is a linear, then $V^*(\vec{x}, t)$ is convex in $\vec{x}$.

Proof: For a wind realization $W(t, \omega)\omega \in \Omega$, and a pair of initial conditions $\vec{x}^{(0)}(0)$ and $\vec{x}^{(1)}(0)$, we take a convex combination of the states $\vec{x}^{(\alpha)}(0) := (1-\alpha)\vec{x}^{(0)}(0) + \alpha\vec{x}(1)$ and provide the convex combination of the optimal power supplied from the former initial states, i.e., $\vec{P}^{(\alpha)} = (1-\alpha)\vec{P}^{(0)} + \alpha\vec{P}^{(1)}$. Due to linearity we obtain the trajectory $\vec{x}^{(\alpha)}(t, \omega) = (1-\alpha)\vec{x}^{(0)}(t, \omega) + \alpha\vec{x}^{(1)}(t, \omega)$. Since $\vec{x}^{(0)}(t, \omega), \vec{x}^{(1)}(t, \omega) \in [\Theta_m, \Theta_M]$, it follows $\vec{x}^{(\alpha)}(t, \omega) \in [\Theta_m, \Theta_M]$. Letting $V^\alpha(x^{(\alpha)}(0))$ be the cost for this allocation, we obtain,

$$(1-\alpha)V^*_\omega(\vec{x}^{(0)}(0)) + \alpha V^*_\omega(\vec{x}^{(1)}(0))$$
$$= \int_0^T (1-\alpha)\mathbf{1}^T(\vec{x}^{(0)}(t) - \Theta_M)^{+2} + \alpha\mathbf{1}^T(\vec{x}^{(1)}(t) - \Theta_M)^{+2} dt$$
$$\geq \int_0^T \mathbf{1}^T[(\vec{x}^{(\alpha)}(t)) - \Theta_M)^+]^2 dt = V^\alpha_\omega(\vec{x}^{(\alpha)}(0)) \geq V^*_\omega(\vec{x}^{(\alpha)}(0)).$$

Taking the expectation proves the desired result. ■

Proof of Theorem 2: Notice that Lemma (1) implies that $\frac{\partial V^*}{\partial x_i} \geq 0$ for i=1,2. Equation (10) thus implies

$$(P_1^*, P_2^*) = \begin{cases} \text{(W,0)} & \text{if } \frac{\partial V_1^*}{\partial x_1} > \frac{\partial V_1^*}{\partial x_2} \\ \text{(0,W)} & \text{if } \frac{\partial V_1^*}{\partial x_1} < \frac{\partial V_1^*}{\partial x_2} \end{cases}.$$

Also, since $V_i^*(\vec{x}, t)$ is convex in $\vec{x}$, we have $\frac{\partial V_i^*(x_1, x_2)}{\partial x_i}$ increases with x_i for $i = 1, 2$. So if (x_1, x_2) is a temperature state with $x_1 > x_2$, then $\frac{\partial V_i^*(\vec{x})}{\partial x_1} > \frac{\partial V_i^*(\vec{x})}{\partial x_2}$, so the minimizer in the HJB equation is $P_1^* = W, P_2^* = 0$. Therefore x_1 decreases while x_2 increases till $x_1 = x_2$. So we conclude that optimal policy is synchronizing in nature. ■

B. HJB equation for hard threshold model

Let $V_i^*(\vec{x}(0))$ be the optimal cost-to-go from $\vec{x}(0)$ when wind state is i (for $i = 1$ when wind is *on*, $i = 0$ when wind is *off*). The control variables are $\vec{P}^g$ and $\vec{P}^w$, the allocated grid and wind power vectors to the TCLs. The state-space is $\mathcal{S} = [\Theta_m, \Theta_M] \times [\Theta_m, \Theta_M] \times \{0, 1\}$. Let $\widehat{\mathcal{A}}(\vec{x}, i) \subset \mathbb{R}^4_+$ be the set of state dependent *admissible* actions.

$$\widehat{\mathcal{A}}(\vec{x}, i) = \{(\vec{P}^g, \vec{P}^w) : P_1^w + P_2^w \leq W \text{ if (Wind } i\text{=1)},$$
$$\text{if (Wind } i\text{=0) } P_1^w = P_2^w = 0,$$
$$\text{if } (x_j(t) = 0) \text{ then } f(P_j^g + P_j^w) \geq 0,$$
$$\text{if } (x_j(t) = \Theta_M) \text{ then } f(P_j^g + P_j^w) \leq 0\}.$$

The Hamilton-Jacobi-Bellman equation is,

$$\inf_{(\vec{P}^g, \vec{P}^w) \in \widehat{\mathcal{A}}(\vec{x}, i)} ((\mathbf{1}^T \vec{P}^g)^2 + \nabla^T V_i^*(\vec{x}, t)\vec{f}(\vec{P}^g + \vec{P}^w)) - q_i V_i^*(\vec{x}, t)$$
$$+ q_i V_{!i}^*(\vec{x}, t)\} + \frac{\partial V_i^*}{\partial t}(\vec{x}, t) = 0 \quad \text{For } i = 0, 1. \quad (11)$$

For homogeneous loads with the state dynamics $f(P) = S - P$ (here S=heating due to ambient temperature when $P = 0$), the infimum term for control separates. So for $x_i \in (0, \Theta_M)$, the optimal wind and grid power allocations are given by

$$\vec{P}^{*g}(\vec{x},i) = \underset{\vec{P}^g\geq 0}{\arg\inf}\left((P_1^g+P_2^g)^2 - \frac{\partial V_i^*}{\partial x_1}P_1^g - \frac{\partial V_i^*}{\partial x_2}P_2^g\right) \quad (12)$$

$$\vec{P}^{*w}(\vec{x}) = \underset{(P_1^w+P_2^w=W)}{\arg\inf}\left(-\frac{\partial V_1^*}{\partial x_1}P_1^w - \frac{\partial V_1^*}{\partial x_2}P_2^w\right). \quad (13)$$

Lemma 3: The optimal cost-to-go function $V^*(\vec{x},t)$ is component wise increasing in $\vec{x}$.

Proof: We use the same construction as Lemma 1. Letting $(\hat{P}_1(t,\omega),P_2(t,\omega))$ be the power used for trajectory $(\hat{y}_1(t,\omega),x_2(t,\omega))$, we have

$$\begin{aligned}V_\omega^*(x_1(0),x_2(0)) &= \int_0^T(\overset{\star}{P}{}_1^g+\overset{\star}{P}{}_2^g)^2ds\\ &\geq \hat{V}_\omega(y_1(0),x_2(0)) \geq V_w^*(y_1(0),x_2(0)).\end{aligned}$$

Taking expectations yields the result. ■

Lemma 4: $V^*(\vec{x},t)$ is convex in $\vec{x}$.

Proof: We use the same construction as Lemma 2. Noting that $\vec{x}^{(\alpha)}(t,\omega)\in[\Theta_m,\Theta_M]$ and $(\vec{P}_g^{(\alpha)},\vec{P}_w^{(\alpha)})\in\widehat{\mathcal{A}}(\vec{x},i)$, we have

$$\begin{aligned}&(1-\alpha)V_\omega^*(\vec{x}^{(0)})+\alpha V_\omega^*(\vec{x}^{(1)})\\ &=(1-\alpha)\int_0^T(\mathbf{1}^T\vec{P_g}^{(0)}(t,\omega))^2+\alpha\int_0^T(\mathbf{1}^T\vec{P_g}^{(1)}(t,\omega))^2ds\\ &\geq\int_0^T(\mathbf{1}^T(1-\alpha)\vec{P_g}^{(0)}(t,\omega)+\alpha\vec{P_g}^{(0)}(t,\omega)))^2ds\\ &=V_\omega^{(\alpha)}(x^{(\alpha)}(0))\geq V_\omega^*(x^{(\alpha)}(0)).\end{aligned}$$

taking expectation establishes the convexity of $V^*(\vec{x})$. ■

Proof of Theorem 3: Equations (12),(13) specify the optimal power allocation. Lemma (3) proves that $\frac{\partial V_i^*}{\partial x_j}\geq 0$, and from Lemma (4), $\frac{\partial V^*}{\partial x_j}(\vec{x})$ is a increasing function of x_j. Therefore when $x_1<x_2$, $\frac{\partial V^*}{\partial x_1}\leq\frac{\partial V^*}{\partial x_2}$. Using this, the minimizers in (12)-(13) are as follows

$$(\overset{\star}{P}{}_1^w,\overset{\star}{P}{}_2^w)=\begin{cases}\text{(W,0)} & \text{if } \frac{\partial V_1^*}{\partial x_1}>\frac{\partial V_1^*}{\partial x_2}.\\ \text{(0,W)} & \text{if } \frac{\partial V_1^*}{\partial x_1}<\frac{\partial V_1^*}{\partial x_2}.\end{cases}. \quad (14)$$

$$(\overset{\star}{P}{}_1^g(\vec{x},i),\overset{\star}{P}{}_2^g(\vec{x},i))=\begin{cases}(\frac{1}{2}\frac{\partial V_i^*}{\partial x_1}(\vec{x}),0) & \text{if } \frac{\partial V_i^*}{\partial x_1}>\frac{\partial V_i^*}{\partial x_2}\\ (0,\frac{1}{2}\frac{\partial V_i^*}{\partial x_2}(\vec{x})) & \text{if } \frac{\partial V_i^*}{\partial x_1}<\frac{\partial V_i^*}{\partial x_2}\\ (\frac{1}{2}\frac{\partial V_i^*}{\partial x_1}(\vec{x}),\frac{1}{2}\frac{\partial V_i^*}{\partial x_2}(\vec{x})) & \text{if } \frac{\partial V_i^*}{\partial x_1}=\frac{\partial V_i^*}{\partial x_2}.\end{cases} \quad (15)$$

When wind is *off*, the power allocation ($\overset{\star}{P}{}^g$) is the same as (15) with V_1^* replaced by V_0^*. Thus we see that the optimal grid power is allocated so that only the higher temperature TCL is cooled, until the lower temperature TCL increases to the its temperature. Thereafter, both the TCLs are cooled at the same rate. Therefore we conclude the optimal policy is of *synchronizing* nature. ■

C. Non-synchronized response under stochastic user preferences

The HJB equation in this case is the same as in the hard threshold model, but the admissible actions need to be modified when the TCLs are above the thermostat set-point:

$$\widetilde{\mathcal{A}}(\vec{x},i,Th)=\{(\vec{P}^g,\vec{P}^w)\in\widehat{\mathcal{A}}(\vec{x},i):$$
$$\text{if } (x_j(t))>\Theta_{max}^1 \text{ and } Th(t)=\Theta_{max}^1) \text{ then } P_j^g=M\}.$$

Lemma 5: If the optimal cost-to-go function is concave in the interval $[a,b]\times[a,b]\subset[0,\Theta_M]\times[0,\Theta_M]$, then the optimal policy is *de-synchronizing* in nature in $[a,b]\times[a,b]$.

Proof: Notice that the HJB equations are the same as (11), with $\widehat{\mathcal{A}}$ replaced by $\widetilde{\mathcal{A}}$. When Th(t) is Θ_M^1 all the TCLs having temperature above Θ_M^1 need to be necessarily provided grid power M. This is neither *synchronizing* nor *de-synchronizing*. Otherwise, the solution is given by equations (14) and (15).Now we observe that when the optimal cost-to-go is locally concave $\frac{\partial V_i^*(\vec{x})}{\partial x_1}\geq\frac{\partial V_i^*(\vec{x})}{\partial x_2}$, when $x_1<x_2$, thus both grid and wind power are allocated to TCL1, thereby making it even cooler, while at the same time TCL2's higher temperature is rising even higher. Therefore the actions taken are *de-synchronizing*. ■

VI. NUMERICAL METHOD FOR COMPUTATION OF SOLUTION

In this section we numerically compute the solution of HJB equation for the *stochastic user preference model*, and show the local concavity of the optimal cost-to-go function. Since the HJB equations (11) are a system of non-linear partial differential equations, an analytical solution is difficult to obtain.To compute the exact solution we use the value-iteration method. To that end, we discretize both the temperature levels and time. We also convert the continuous time Makrov processes for wind and thermostat settings into two-state discrete time Markov chains. To obtain a state space which is the set of integers, we use integer values for M and S, and optimize with respect to $\vec{P}^g,\vec{P}^w$ in the integer action space. This results in the following infinite time discounted cost problem with discount factor $\beta\in(0,1)$:

$$\begin{aligned}\text{Minimize } & \sum_{t=0}^{\infty}\beta^t\mathbb{E}[(\mathbf{1}^T\vec{P}^g[t])^2]\\ \text{Sub. to. } & x_i[t+1]=x_i[t]+S-P_i^g-P_i^w \text{ for } i=1,2\\ & x_1,x_2\in\{0,1,2...,\Theta_M\}\\ & P_1^w+P_2^w=W[t]\\ & W[t]\backsim\mathcal{M}[q_0,q_1],P_i^g\geq 0,P_i^w\geq 0 \text{ for } i=1,2.\end{aligned}$$

We use the value iteration algorithm to obtain $V_i[\vec{x}]$ for $i=1,2$. Figure 1 illustrates the local concavity in optimal cost-to-go function computed numerically. Figure 2 exhibits a simulation under the optimal policy, for the initial conditions $x_1(0)=x_2(0)=95$. It shows that the optimal policy de-synchronizes the two TCLs when either of the TCLs is at high temperature, while it attempts to synchronize them when they are at low temperatures.

VII. HEURISTIC APPROXIMATION

We now propose a simple heuristic approximation of the optimal policy which attempts to capture its main characteristics. To do so, it is useful to plot the vector field of rate of change of the temperature state vector. Figure 3 indicates a threshold above which the TCLs are de-synchronized; let us denote such a level by $\tau_h^{(i,j)}$ where i,j denote the wind condition and thermostat setting respectively. Similarly we can approximate the magnitude of the grid power allocated by two linear splines, one above $\tau^{(i,j)}$ and one below, as shown in Figure 4. Therefore the overall heuristic policy has five parameters $(\tau_1,\tau_2,P_{h1},P_{l1},P_{h2})$ per state. Figure 5

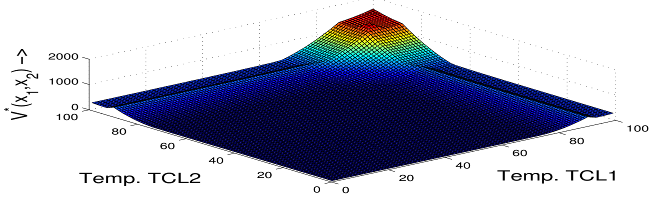

Fig. 1. The optimal cost-to-go function is concave in the higher temperature region, but is convex in the lower temperature region.

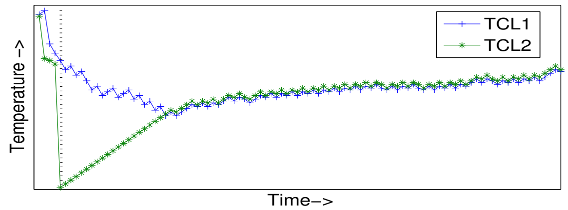

Fig. 2. Simulation under the optimal policy. Note that when either of the TCLs is above a threshold temperature, the TCLs are separated apart, while at low temperatures, the TCLs are brought to the same temperature.

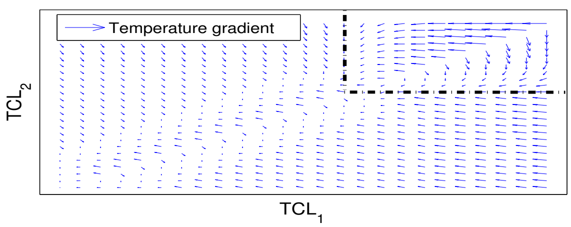

Fig. 3. Vector field of the rate of change of temperature states under the optimal policy when wind is *off* and thermostat is *high*. At higher temperatures the TCLs are separated, while at low temperature they are brought closer together.

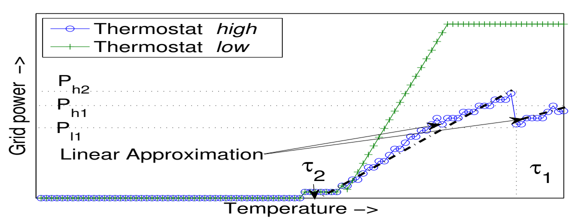

Fig. 4. Magnitude of the total grid drawn by $N = 2$ TCLs when wind is *on*, under the numerically computed optimal policy, when $x_1 = x_2$. Note that it can be linearly approximated by parameters $(P_{h1}, P_{h2}, P_{l1}, \tau_1, \tau_2)$.

compares the grid power drawn by the heuristic policy $\hat{\Pi}$, the numerically computed optimal synchronizing policy Π_{sync}, and a passive bang-bang policy Π_{bang} for which the TCLs are turned on and off when they hit Θ_m and Θ_M respectively. One can note that the variations in power trajectory are lesser under the heuristic policy than under the synchronized policy.

VIII. CONCLUSION & FUTURE WORK

We have studied the nature of the optimal demand response for the stochastic control problem of renewable power allocation to thermal inertial loads. Starting with a very simple model, we have incorporated several considerations

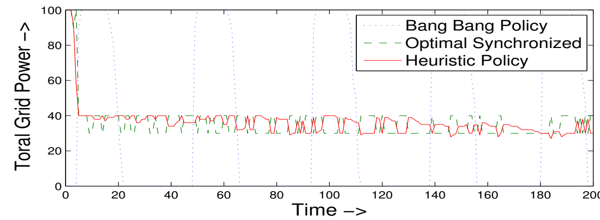

Fig. 5. Comparison of grid power drawn for different policies for $N = 10$, $\Theta_M^2 = 100, \Theta_M^1 = 70$. The heuristic de-synchronizing policy is smoother than the synchronizing policy.

for the optimizing schemes. We have provided a theoretical foundation for the staggered nature of such optimizing demand responses. Our analysis is mostly focused on a simple system illustrating a homogeneous population of TCLs. A comprehensive treatment along with an efficient computational algorithm will potentially be a useful next step.

REFERENCES

[1] S. Kundu and N. Sinitsyn, "Safe protocol for controlling power consumption by a heterogeneous population of loads," *American Control Conference*, pp. 2947–2952, 2012.

[2] S. Kundu, N. Sinitsyn, S. Backhaus, and I. Hiskens, "Modelling and control of thermostatically controlled loads," in *Proc. 17th Power Systems Computation Conference*, 2011.

[3] D. Callaway, "Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy," *Energy Conversion and Management*, vol. 50, no. 9, pp. 1389–1400, 2009.

[4] Z. Xu, J. Ostergaard, M. Togeby, and C. Marcus-Moller, "Design and modelling of thermostatically controlled loads as frequency controlled reserve," *Power Engineering Society General Meeting*, pp. 1–6, 2007.

[5] J. A. Short, D.G.Infield, and L. Freris, "Stabilization of grid frequency through dynamic demand control," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1284–1293, 2007.

[6] A. Molina-Garca, M. Kessler, J. A. Fuentes, and E. Gmez-Lzaro, "Probabilistic characterization of thermostatically controlled loads to model the impact of demand response programs," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 241–251, 2011.

[7] N. Lu and D. P. Chassin, "A state-queueing model of thermostatically controlled appliances," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1666–1673, 2004.

[8] M. D. Ilic, L. Xie, and J.-Y. Joo, "Efficient coordination of wind power and price responsive demand-Part I: Theoretical foundations," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 1875–1884, 2011.

[9] P. Anthony and S. S. Oren, "Supplying renewable energy to deferrable loads: Algorithms and economic analysis," *IEEE Power and Energy Society General Meeting*, pp. 1–8, 2010.

[10] N. Lu and S. Katipamula, "Control strategies of thermostatically controlled appliances in a competitive electricity market," *IEEE Power and Energy Society General Meeting,*, vol. 1, pp. 202–207, 2005.

[11] P. Pinson, L. Christensen, H. Madsen, P. Sorensen, M. Donovan, and L. Jensen, "Regime-switching modelling of the uctuations of oshore wind generation," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 96, no. 12, pp. 2327–2347, 2008.

[12] R. Billinton, H. Chen, and R. Ghajar, "Time-series models for reliability evaluation of power systems including wind energy," *Microelectronics Reliability*, vol. 36, no. 9, pp. 1253–1261, 1996.

[13] L. Kroldrup. (2010) Gains in global wind capacity. [Online]. Available: http://green.blogs.nytimes.com/2010/02/15/gains-in-global-wind-capacity-reported/

[14] R. Pernick, "Clean energy trends 2012," Clean Edge, Tech. Rep., 2012.

[15] (2011) Global status report. [Online]. Available: http://new.ren21.net/

[16] J. Mathieu, S. Koch, and D. Callaway, "State estimation and control of electric loads to manage real-time energy imbalance," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 430–440, 2013.