



HAL
open science

Kullback-Leibler-Quadratic Optimal Control of Flexible Power Demand

Neil Cammardella, Ana Bušić, Yuting Ji, Sean Meyn

► **To cite this version:**

Neil Cammardella, Ana Bušić, Yuting Ji, Sean Meyn. Kullback-Leibler-Quadratic Optimal Control of Flexible Power Demand. CDC 2019 - 58th IEEE Conference on Decision and Control, Dec 2019, Nice, France. hal-02425974

HAL Id: hal-02425974

<https://hal.science/hal-02425974>

Submitted on 31 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Kullback-Leibler-Quadratic Optimal Control of Flexible Power Demand

Neil Cammardella¹, Ana Bušić², Yuting Ji³, and Sean Meyn⁴

Abstract—A new stochastic control methodology is introduced for distributed control, motivated by the goal of creating virtual energy storage from flexible electric loads, i.e. *Demand Dispatch*. In recent work, the authors have introduced Kullback-Leibler-Quadratic (KLQ) optimal control as a stochastic control methodology for Markovian models. This paper develops KLQ theory and demonstrates its applicability to demand dispatch. In one formulation of the design, the grid balancing authority simply broadcasts the desired tracking signal, and the heterogeneous population of loads ramps power consumption up and down to accurately track the signal. Analysis of the Lagrangian dual of the KLQ optimization problem leads to a menu of solution options, and expressions of the gradient and Hessian suitable for Monte-Carlo-based optimization. Numerical results illustrate these theoretical results.

I. INTRODUCTION

The replacement of controllable fossil-fuel generators by volatile renewable generation is creating new challenges for power grid operators, who must continuously match demand with supply to maintain reliable operation. Consequently, some have begun spending millions of dollars on grid-scale batteries. Do we require this massive infrastructure investment? Research demonstrates that distributed control of flexible loads can provide battery-like services such as frequency regulation while guaranteeing quality of service (QoS) to the consumer [1], [2], [3], [4]. This emerging science is called *Demand Dispatch*.

The present paper is motivated by [5], in which a net load forecast (total power demand minus wind and solar) is used to calculate an optimal schedule for flexible loads. Dynamic models for thermostatically-controlled loads [6], [7] and other loads are used to enforce constraints on power and energy, ensuring QoS to the consumers. Fig. 1 shows the results of one experiment from [5]. Demand dispatch is able to provide many GWs of flexibility, which allows non-renewable generation to follow a much flatter trajectory, and potentially raises the ceiling on renewable energy penetration.

Funding provided through the ARPA-E NODES award, *Open and Scalable Distributed Energy Resource Networks*, National Science Foundation awards EPCN 1609131 & CPS 1646229, and French National Research Agency grant ANR-16-CE05-0008.

¹Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611

²Inria and DI ENS, École Normale Supérieure, CNRS, PSL Research University, Paris, France

³Civil and Environmental Engineering Department, Stanford University, Stanford, CA 94305

⁴Department of Electrical and Computer Engineering, University of Florida, and Inria International Chair, Paris

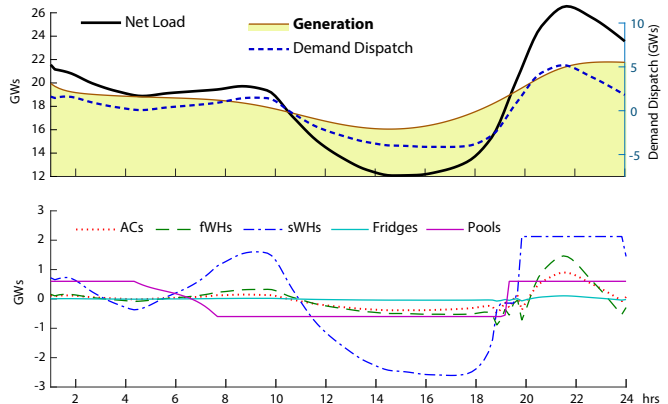


Fig. 1: Taming the Duck Curve with feed-forward control

A salient question, only partially addressed in [5], remains unanswered: how is an optimal demand dispatch schedule realized at full capacity, without centralized control? In this paper, we answer this question by posing the problem as optimal control of a Markov decision process at each load. To balance the needs of the grid and the consumers, we propose a control objective that is a sum of two terms: a quadratic penalty for tracking error plus the Kullback-Leibler (K-L) divergence of the controlled stochastic model with respect to the nominal stochastic model. We show a unique solution exists and present convergent algorithms to compute it.

A. Control Architecture Overview

Our goal is to modify the behavior of loads so that their aggregate power consumption tracks a reference signal $r = \{r_k\}$ that is broadcast by a *Balancing Authority* (BA). An individual load is modeled as a Markov chain \mathbf{X} , with finite state space $\mathbf{X} = \{x^1, \dots, x^d\}$, evolving on a discrete time horizon $0, 1, \dots, M$. The nominal stochastic model is defined by a probability mass function (pmf) p^0 that defines the nominal pmf of (X_0, \dots, X_M) . We use \vec{x} to denote a generic element of \mathbf{X}^{M+1} and $\mathcal{P}(\mathbf{X}^{M+1})$ to denote the space of pmfs on \mathbf{X}^{M+1} . The power consumption of the load in state x_k is $\mathcal{Y}(x_k)$. The marginal pmfs are denoted $\{\nu_k^0 : 0 \leq k \leq M\}$ and the transition matrices $\{P_k : 0 \leq k \leq M-1\}$ so that the nominal pmf can be factored as:

$$p^0(\vec{x}) = \nu_0^0(x_0)P_0(x_0, x_1)P_1(x_1, x_2) \cdots \quad (1)$$

To achieve our goal, we construct a Kullback-Leibler-Quadratic (KLQ) optimal control objective; fix $\kappa > 0$, and

minimize over p :

$$J(p; \nu_0^0) := D(p||p^0) + \frac{\kappa}{2} \sum_{k=1}^M [\langle \nu_k, \mathcal{Y} \rangle - r_k]^2 \quad (2)$$

s.t. $\nu_0 = \nu_0^0$

where ν_0 is the marginal of p at time 0. The relative entropy is the mean log-likelihood:

$$D(p||p^0) = \sum_{\vec{x}} L(\vec{x}) p(\vec{x}), \quad L = \log(p/p^0). \quad (3)$$

It is shown in this paper that the optimizing pmf on \mathcal{X}^{M+1} can be expressed as a tilting of the nominal pmf, of the form:

$$p^\beta(\vec{x}) = p^0(\vec{x}) \exp(L_\beta(\vec{x})) \quad (4)$$

in which the log-likelihood ratio is

$$L_\beta(\vec{x}) = \mathcal{Y}_\beta(\vec{x}) - \Lambda_\beta(x_0), \quad (5)$$

where $\mathcal{Y}_\beta(\vec{x}) = \sum_{k=1}^M \beta_k \mathcal{Y}(x_k)$, $\beta \in \mathbb{R}^M$ are Lagrange multipliers, and $\Lambda_\beta(x_0)$ is viewed as a normalizing constant for each $x_0 \in \mathcal{X}$. It follows that $p^\beta(\vec{x})$ is Markovian and can be factored in a form similar to (1) with transition matrices $\{P_k^\beta : 0 \leq k \leq M-1\}$ and marginal pmfs $\{\nu_k^\beta : 0 \leq k \leq M\}$; henceforth, we suppress the superscript in ν_k^β to lighten the notation:

$$p^\beta(\vec{x}) = \nu_0(x_0) P_0^\beta(x_0, x_1) P_1^\beta(x_1, x_2) \cdots \quad (6)$$

Mean-field theory is used to model aggregate load behavior. For N loads, the empirical pmf at time k is defined as

$$\nu_k^{(N)}(x) := \frac{1}{N} \sum_{i=1}^N \mathbb{I}\{X_k^i = x\}, \quad x \in \mathcal{X}. \quad (7)$$

It can be shown that the approximation $\nu_k^{(N)} \approx \nu_k$ becomes exact as $N \rightarrow \infty$. The mean-field model is the deterministic system defined by

$$\nu_{k+1} = \nu_k P_k^\beta \quad (8)$$

in which ν_k is interpreted as a d -dimensional row vector, and P_k^β a $d \times d$ matrix for each k .

An example of mean field dynamics is illustrated in Fig. 2. This shows the evolution of the $\{\nu_k\}$ for a particular example described in Section IV. Although the initial pmfs are very different, they become nearly identical after three hours. This suggests that $\{\nu_k\}$ can be estimated in a distributed manner, which has important implications for distributed control. Further discussion is contained in Section IV.

B. Literature Review

The goal in much of the prior work [8], [9], [13], [14], [2], [15] and in the present paper is to modify the behavior of loads so that their aggregate power consumption tracks a reference signal. Control design is based on a mean-field model where p^0 is an approximation of the behavior of the collection of loads in the absence of load control. In the context of power systems applications, mean-field models

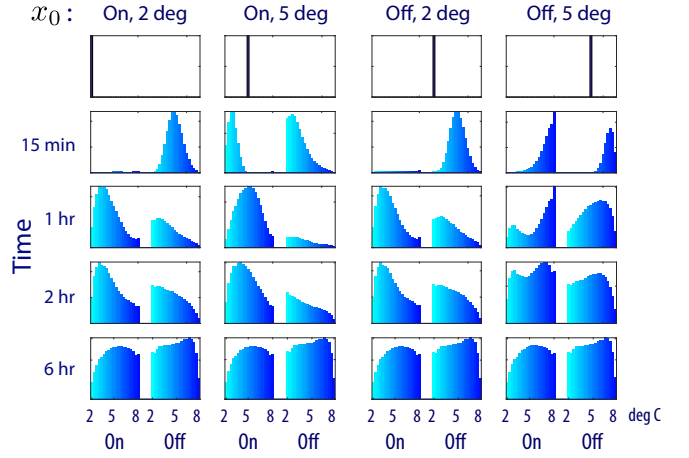


Fig. 2: Evolution of the marginals $\{\nu_k^*\}$ for a refrigerator model from four different initial (degenerate) marginals.

first appeared in [10], [11]. Conditions are imposed there and in later work [12], [13], [2] so that the mean-field approximation becomes exact as the number of loads approaches infinity. Randomization is introduced for two reasons: first, it leads to a convex relaxation of a combinatorial online decision problem; second, just as in telecommunications, randomization is a simple means to avoid synchronization of the population [1], [14], [15].

The current paper introduces new distributed control techniques based on finite horizon optimal control, building on the individual perspective design of [3]. It is related to [16], [17], [18], in that relative entropy is used as a penalty term to discourage deviations from nominal behavior. A similar optimal control approach is briefly described in Section 5 of the book chapter [19]:

$$\begin{aligned} \min_p D(p||p^0) \\ \text{subject to } E_p[\mathcal{Y}(X_k)] = r_k. \end{aligned} \quad (9)$$

The optimization criterion (9) is precisely one-step dead-beat control: the constraint $y_k \equiv r_k$, with $y_k = E_p[\mathcal{Y}(X_k)]$, means that we wish to partially invert the system dynamics. The potential challenge is most clear when we realize that we eventually pose the problem as a single-input single-output nonlinear system, with input β and output \mathbf{y} , and then the goal is to invert the system to obtain β^* as a function of \mathbf{r} . It is known that that this inversion process can lead to high gain and even instability. This is most easily seen here by considering a small gain analysis: if \mathbf{r} and β are small, then the dynamics from β to \mathbf{y} can be linearized, and high gain is a consequence of inverting the linear mapping between these two signals. This solution may be highly sensitive to model error. The KLQ setting allows us to make the right tradeoff between tracking and gain, and learn what reference signals are easily followed by the collection of loads.

For the KLQ control objective proposed in this paper, the tracking constraint in (9) is replaced by a quadratic loss function and a constraint is added to keep the initial distribution unchanged. This provides many advantages:

- (i) The approach is less sensitive to modeling error

- (ii) An optimal solution always exists, while feasibility of (9) requires conditions on p^0 and r
- (iii) Tradeoffs can be made between tracking performance and deviating from nominal behavior
- (iv) The approach respects the physical impossibility of changing the initial distribution, while (9) only respects this constraint for degenerate distributions.

As $\kappa \rightarrow \infty$ we recover the solution to the dead-beat control problem (9) only if ν_0^0 is degenerate, so that $\nu_0 = \nu_0^0$ is automatic when $D(p||p^0) < \infty$.

The rest of the paper is organized as follows: Section II contains a proof of the existence of a unique optimal solution to the KLQ problem. A Lagrangian dual formulation is presented in Section III, leading to convergent algorithms. Section IV contains numerical results illustrating KLQ solutions, and further insight on how this can be applied to demand dispatch in practice. The section contains discussion of possible control architectures. Section V concludes the paper and contains directions for future research.

II. KLQ OPTIMALITY

We begin with a characterization of optimality:

Theorem 2.1. *Consider the problem of minimizing $J(p; \nu_0^0)$ over the set of pmfs p with given initial marginal ν_0^0 . An optimizer p^* exists and is unique. The log-likelihood $L^* = \log(p^*/p^0)$ is given by*

$$L^*(\vec{x}) = \mathcal{Y}_{\beta^*}(\vec{x}) - \Lambda_{\beta^*}(x_0), \quad (10)$$

The vector $\beta^* \in \mathbb{R}^M$ and function $\Lambda^*: \mathcal{X} \rightarrow \mathbb{R}$ solve

$$\beta_k^* = \kappa [r_k - \langle \nu_k^*, \mathcal{Y} \rangle], \quad 1 \leq k \leq M \quad (11)$$

$$\Lambda_{\beta^*}(x_0) = \log(\mathbb{E}_{p^0}[\exp(\mathcal{Y}_{\beta^*}(\vec{x})) \mid X_0 = x_0]). \quad (12)$$

Proof. Existence of a unique solution to (11), (12) is contained in Prop. 3.1. Optimality is established in Lemma 2.3; the proof is based on Kullback's Lemma [20], which is summarized in Lemma 2.2. \square

In this finite state-space setting, (12) becomes

$$\Lambda_{\beta^*}(x_0) = \log \sum_{x_1, \dots, x_M} p^0(\vec{x} \mid x_0) \exp\left(\sum_{k=1}^M \beta_k^* \mathcal{Y}(x_k)\right),$$

where the conditional pmf is defined via Bayes' rule:

$$p^0(x_0, \dots, x_M \mid x_0) = \frac{1}{\nu_0^0(x_0)} p^0(x_0, x_1, \dots, x_M)$$

Thm. 2.1 follows from Kullback's Lemma [20]:

Lemma 2.2. *For $p^0, p^1, p^2 \in \mathcal{P}(\mathcal{X}^{M+1})$:*

$$D(p^2||p^0) \geq D(p^1||p^0) + \langle p^2 - p^1, L \rangle$$

in which $L = \log(p^1/p^0)$. \square

The next lemma establishes Thm. 2.1 under the assumption that there exists a solution β^* to (11). The fact that a solution always exists is postponed to Prop. 3.1.

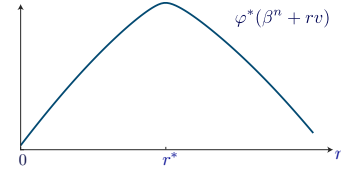


Fig. 3: A plot of the dual function along a line-segment

Lemma 2.3. *Suppose that there exists a solution to the fixed point equation (11). Then p^* , with log likelihood ratio given in (10), is the optimizer of $J(p; \nu_0^0)$ over all pmfs p with initial marginal ν_0^0 .*

Proof. We begin with an application of Lemma 2.2 using $p^2 = p$ and $p^1 = p^*$. The quadratic term in $J(p; \nu_0^0)$ admits a lower bound, starting with the simple inequality $x^2 \geq y^2 + 2y(x - y)$. Applying this with $x = \langle \nu_k, \mathcal{Y} \rangle - r_k$ and $y = \langle \nu_k^*, \mathcal{Y} \rangle - r_k$ gives

$$\begin{aligned} [\langle \nu_k, \mathcal{Y} \rangle - r_k]^2 &\geq [\langle \nu_k^*, \mathcal{Y} \rangle - r_k]^2 \\ &\quad + 2[\langle \nu_k^*, \mathcal{Y} \rangle - r_k] \langle \nu_k - \nu_k^*, \mathcal{Y} \rangle \end{aligned}$$

We have $\langle \nu_k^*, \mathcal{Y} \rangle - r_k = -\beta_k^*/\kappa$ under the assumptions of the lemma. Combining the inequality with Lemma 2.2:

$$\begin{aligned} J(p; \nu_0^0) &\geq D(p^*||p^0) + \langle p - p^*, L^* \rangle \\ &\quad + \frac{\kappa}{2} \sum_{k=1}^M [\langle \nu_k, \mathcal{Y} \rangle - r_k]^2 \\ &\geq J(p^*; \nu_0^0) + \langle p - p^*, L^* \rangle \\ &\quad + \kappa \sum_{k=1}^M (-\beta_k^*/\kappa) \langle \nu_k - \nu_k^*, \mathcal{Y} \rangle \end{aligned} \quad (13)$$

Since $\nu_0 = \nu_0^0$, we have $\langle p - p^*, L^* \rangle = \langle p - p^*, \mathcal{Y}_{\beta^*} \rangle$:

$$\begin{aligned} J(p; \nu_0^0) &\geq J(p^*; \nu_0^0) + \langle p - p^*, \mathcal{Y}_{\beta^*} \rangle \\ &\quad - \sum_{k=1}^M \beta_k^* \langle \nu_k - \nu_k^*, \mathcal{Y} \rangle \\ &= J(p^*; \nu_0^0) \end{aligned} \quad (14)$$

where the final equality follows from the definition of \mathcal{Y}_{β^*} . \square

III. DUALITY AND ALGORITHMS

The difficulty with the fixed-point formula (11) is that it is not obvious that standard algorithms such as successive approximation will find a solution. In Prop. 3.1 that follows, the fixed point equation is interpreted as a first-order condition for optimality of a concave function. Based on this result we are led to existence of a solution, and convergent algorithms to compute β^* .

The minimization of $J(p; \nu_0^0)$ is expressed as the following convex program:

$$\begin{aligned} J^*(\nu_0^0) &:= \min_{p, \gamma} D(p||p^0) + \frac{\kappa}{2} \sum_{k=1}^M \gamma_k^2 \\ \text{s.t. } &\gamma_k = \langle \nu_k, \mathcal{Y} \rangle - r_k, \quad 1 \leq k \leq M \\ &\nu_0 = \nu_0^0 \end{aligned} \quad (15)$$

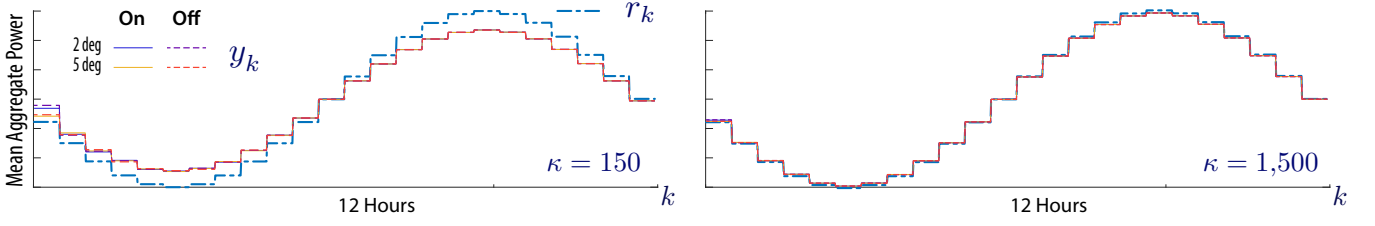


Fig. 4: Tracking for a refrigerator model from four different initial conditions, with two different values of κ . The aggregate power consumption nearly coincides after about three hours with $\kappa = 150$, and coupling occurs much faster when κ is increased to 1,500.

Let $\beta \in \mathbb{R}^M$ denote a Lagrange multiplier for the first constraint and h the Lagrange multiplier for the marginal constraint. While it is most natural to regard h as a real-valued function on X , for computational simplicity we take $h: \mathsf{X}^{M+1} \rightarrow \mathbb{R}$, with the understanding that it only depends on x_0 .

The Lagrangian is a function of the variables p, γ, β, h :

$$\begin{aligned} \mathcal{L}(p, \gamma, \beta, h) = & D(p \| p^0) + \frac{\kappa}{2} \sum_{k=1}^M \gamma_k^2 + \langle p - p^0, h \rangle \\ & + \sum_{k=1}^M \beta_k [\gamma_k + r_k - \langle \nu_k, \mathcal{Y} \rangle] \end{aligned} \quad (16)$$

The dual function is defined as the infimum

$$\varphi^*(\beta, h) = \inf_{\gamma, p} \mathcal{L}(p, \gamma, \beta, h)$$

Strong duality holds: $\max_{\beta, h} \varphi^*(\beta, h) = J^*(\nu_0^0)$. Through analysis of this dual convex program we obtain algorithms to obtain the optimizer (β^*, h^*) for the dual, and thence the optimizer (p^*, γ^*) for the primal.

Properties of the dual function are contained in Prop. 3.1. The first order condition for optimality of the Lagrange multipliers corresponding to the average power constraints is shown in (20). It coincides with the fixed point equation for β^* in (11). Consequently, β^* exists and coincides with these optimal Lagrange multipliers. Furthermore, from (18) and (5) it follows that $h^* = \Lambda_{\beta^*}$. A proof is found in Appendix A.

Proposition 3.1. *The dual function and optimizers (β^*, h^*) have the following representations:*

(i) For any $\beta \in \mathbb{R}^M$ and function h , the dual function is

$$\varphi^*(\beta, h) = -\frac{1}{2\kappa} \|\beta\|^2 + \beta^T r - \log \langle p^0, e^{\mathcal{Y}^{\beta-h}} \rangle - \langle p^0, h \rangle \quad (17)$$

(ii) For a given $\beta \in \mathbb{R}^M$, a maximizer of (17) over h is

$$h_{\beta}^*(x_0) = \log \left(\sum_{x_1, \dots, x_M} p^0(\vec{x} | x_0) e^{\mathcal{Y}^{\beta}(\vec{x})} \right) \quad (18)$$

with corresponding value and first derivatives:

$$\begin{aligned} \varphi^*(\beta) & := \max_h \varphi^*(\beta, h) = -\frac{1}{2\kappa} \|\beta\|^2 + \beta^T r - \langle p^0, h_{\beta}^* \rangle \\ \frac{\partial}{\partial \beta_k} \varphi^*(\beta) & = -\frac{1}{\kappa} \beta_k + r_k - \langle \nu_k, \mathcal{Y} \rangle, \quad 0 \leq k \leq M \end{aligned} \quad (19)$$

(iii) The function $\varphi^*(\beta)$ in (19) is strictly concave, and admits a unique maximizer β^* that satisfies the first-order optimality condition; for $0 \leq k \leq M$,

$$0 = \frac{\partial}{\partial \beta_k} \varphi^*(\beta^*) = -\frac{1}{\kappa} \beta_k^* + r_k - \langle \nu_k^*, \mathcal{Y} \rangle \quad (20)$$

□

The remainder of the section contains application of Prop. 3.1 to algorithm design. In algorithms such as steepest ascent to obtain β^* , we construct a sequence of approximating pmfs, each of the form (4). The next result asserts that these pmfs are Markov. A proof is found in Appendix B

Proposition 3.2. *If the nominal model is Markov, then p^{β} is also Markov for any β :*

$$p^{\beta}(\vec{x}) = \nu_0^0(x_0) P_0^{\beta}(x_0, x_1) \cdots P_{M-1}^{\beta}(x_{M-1}, x_M). \quad (21)$$

The transition matrices are constructed via the following:

(i) for each k , a positive matrix is defined:

$$\widehat{P}_k(x, x') = P_k(x, x') \exp(\beta_{k+1} \mathcal{Y}(x')), \quad x, x' \in \mathsf{X}$$

(ii) A sequence of functions $\{a_k: \mathsf{X} \rightarrow \mathbb{R}_+\}$ is defined inductively as follows: $a_M(x) = 1$ for all x and

$$a_k = \widehat{P}_k a_{k+1}, \quad 0 \leq k \leq M-1$$

Then, for each k ,

$$P_k^{\beta}(x, x') = \frac{1}{a_k(x)} \widehat{P}_k(x, x') a_{k+1}(x'), \quad x, x' \in \mathsf{X}$$

□

The steepest ascent algorithm to maximize φ^* , and thereby obtain β^* , is based on the gradient representation in (19). Lemma 3.3 that follows provides representations of the Hessian in terms of the second-order statistics of the random variables $\{\mathcal{Y}(X_k) : 1 \leq k \leq M\}$:

$$\begin{aligned} \Sigma_{i,j}(\beta) & = \mathbb{E}[\mathcal{Y}(X_j) \mathcal{Y}(X_k)] \\ & \quad - \mathbb{E}[\mathbb{E}[\mathcal{Y}(X_j) | X_0] \mathbb{E}[\mathcal{Y}(X_k) | X_0]] \end{aligned} \quad (22)$$

in which the expectations are with respect to p^{β} .

Lemma 3.3. *For any $\beta \in \mathbb{R}^M$, the Hessian has entries*

$$\frac{\partial^2}{\partial \beta_i \partial \beta_j} \varphi^*(\beta) = -\frac{1}{\kappa} \mathbb{I}\{i = j\} - \Sigma_{i,j}(\beta)$$

□

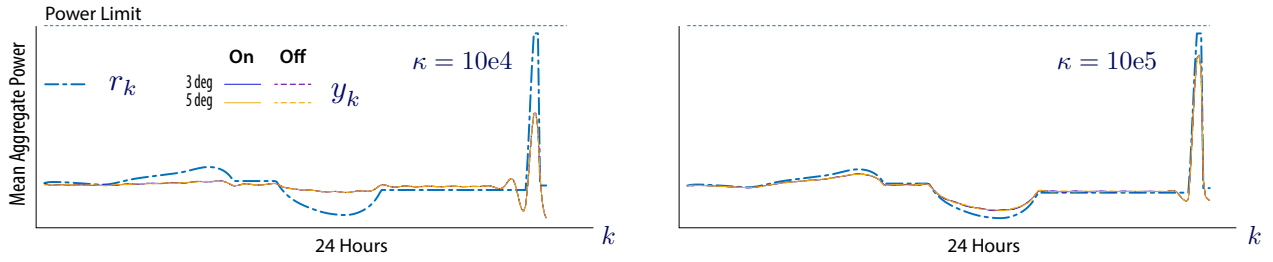


Fig. 5: Tracking of a signal anticipated in California during a very sunny day.

Given the simple form of these derivatives, it is tempting to apply Newton-Raphson to obtain β^* . The difficulty with second order methods is illustrated in Fig. 3. This is a plot of a typical example in which $\beta^n \in \mathbb{R}^M$ is given, $v = \nabla\varphi^*(\beta^n)$, and the plot shows $\varphi^*(\beta^n + rv)$ for a range of positive r . The norm of the Hessian is very nearly zero for $r = r^*$, so that the standard Newton-Raphson method will result in large “overshoot”.

We have also found in examples that steepest ascent may be slow to converge. In the numerical results that follow we opt for line-search using the Golden Section method [21], [22]. Based on Fig. 3, $\beta^{n+1} = \beta^n + r_n^* \nabla\varphi^*(\beta^n)$.

IV. APPLICATION TO DEMAND DISPATCH

Returning to our goal of distributed control of flexible loads, we present numerical results of KLQ control applied to refrigerators. In summary, given a feasible reference signal, tracking becomes excellent with high enough κ . Also, in the scenario where each load optimizes according to its own initial state, rapid coupling of the pmfs is observed, which has interesting implications for distributed control.

Additional structure, described in [3], is imposed on the Markov chain \mathbf{X} . The full state space is the Cartesian product of two finite state spaces: $\mathbf{X} = \mathbf{X}_n \times \mathbf{X}_u$, and we write $\mathbf{X} = (\mathbf{X}_u, \mathbf{X}_n)$: the components \mathbf{X}_u can be directly controlled, and \mathbf{X}_n represents remaining states (such as the temperature evolution within a refrigerator). A deterministic model for \mathbf{X}_n is assumed so that the KLQ control solution does not modify these physical dynamics. The extension of KLQ for fully stochastic models is a topic of current research; see [23], [24], [25] for preliminary results.

A. Coupling in KLQ

Fig. 4 displays the results of two tracking experiments for a refrigerator model from four different initial conditions, with two different values of κ , where $y_k = \langle \nu_k^*, \mathcal{Y} \rangle$. Tracking is nearly perfect with the larger κ .

Observe that the power trajectories from different initial conditions couple rapidly for either value of κ . Fig. 2 shows the controlled marginals becoming very similar within three hours. It is argued in the next section that this can simplify distributed control implementation.

Fig. 5 shows tracking results based on a decomposition of net load in California: the resource allocation algorithm [5] was applied to decompose the desired reference signal as a sum of several signals, distinguished by intended load

class. The algorithm is a quadratic program, in which the objective is a flatter and smoother net load. Power and energy constraints ensure feasibility of the reference signal for each class of loads. The reference signal shown is the power trajectory intended to be tracked by residential refrigerators.

Tracking is excellent for large κ . It is not perfect, but there is good reason for imperfection: the deep belly in the middle of the day is tough for the loads, as it requires the average temperature of the population to stray far from its mean. The spike at the end of the day is somewhat uncomfortable for them, but not so bad. The plots suggest that the cheapest way to track the spike is to introduce a bit of undershoot.

Coupling of aggregate power consumption occurs quickly: on this 24-hour time window, the power consumption from different initial conditions is indistinguishable.

B. Control Architectures

Below are three examples of communication and computation architectures to apply the KLQ method in demand dispatch applications:

- (i) *Smart BA*: The BA uses the reference signal r and its estimate of ν_0^0 to compute β^* and broadcast it to the loads.
- (ii) *Smart Load*: The BA broadcasts r to the loads. Each load computes β^* based on its internal model and $\nu_0^0 = \delta_{x_0}$, with $x_0 \in \mathbf{X}$ its current state.
- (iii) *Genius Load*: The BA broadcasts r to the loads. Each load computes β^* based on its internal model and its estimate of ν_0^0 .

Each approach has advantages and disadvantages. If the BA broadcasts β^* , then it must partition the load population into approximately homogeneous subpopulations and optimize over each. It must also estimate ν_0^0 . If the BA broadcasts r , then it must be tailored for each load, as in [5], or suitably transformed at each load.

The total cost in cases (i) or (iii) is equal to $J^*(\nu_0^0)$ (assuming perfect estimates of ν_0^0). The total cost in the control architecture (ii) is the sum,

$$\sum_{i=1}^d \nu_0(x^i) J^*(\delta_{x^i})$$

since in this case, a load starting at state x^i solves the KLQ optimization problem using the degenerate initial pmf. However, even when the aggregate can easily track r , the cost $J^*(\delta_{x^i})$ may be very large for loads with initial condition x^i that are at odds with the reference signal. For example,

power reduction could be requested while a fridge is near its upper temperature limit and must turn on. So, optimizing with respect to the initial distribution of the population, (i) and (iii), may create less stress on the loads, as compared to (ii).

A significant advantage of approaches (ii) and (iii) is that they automatically address the challenge of heterogeneity of the population. In approach (i) the BA computes a different signal β^* for each load class, based on the respective model. In the other approaches, the load performs calculations based on local observations of its behavior (from which the nominal model is easily constructed locally).

Approach (iii) may appear overly ambitious – how can a load estimate the histogram of the population? We do not expect exactly that. Rather, an implementation of (iii) may be performed as follows: fix a review period T_0 . At a time t_0 at which the KLQ optimization problem is solved, the estimate of ν_{t_0} is defined as the mean-field histogram based on the state of the particular load T_0 time units prior. A glance at Fig. 2 suggests that this may be a good estimate of the population, provided T_0 is chosen with care. There can of course be refinements of this “open loop” approach by applying standard techniques for state estimation/smoothing.

V. CONCLUSIONS

A new stochastic, distributed control technique is presented that uses a quadratic performance penalty plus Kullback-Leibler divergence to balance control actions and deviations from nominal behavior. We show a unique solution exists and propose convergent algorithms based on a Lagrangian dual formulation. Numerical results of KLQ control applied to refrigerators exhibit excellent signal tracking and rapid coupling of the pmfs. Many extensions will be considered in future research.

(i) A deeper understanding of the relationship between optimality and coupling of the pmfs is desired.

(ii) It is assumed for simplicity in this paper that $\mathcal{Y}(x_k)$ is power consumption associated with x_k . In some applications it may be useful to allow this to be a function of two variables, and also time dependent, resulting in

$$\hat{P}_k(x, x') = P_k(x, x') \exp(\beta_{k+1} \mathcal{Y}_k(x, x')), \quad x, x' \in \mathcal{X}$$

The main results go through with only notational changes.

(iii) In an MPC setting, we may view $\mathcal{Y}_M(x_M)$ as a terminal cost function. Careful design of this may result in better performance with smaller time horizon [26].

(iv) We are considering wavelet transforms to reduce complexity, which will be critical when we move to models with continuous state and/or continuous time.

(v) Application of KLQ to simultaneous resource allocation and control is another topic of current research.

APPENDIX

A. Proof of Prop. 3.1

The proof of Prop. 3.1 is based on another well known result regarding relative entropy. For any function $g: \mathcal{X}^{M+1} \rightarrow$

\mathbb{R} denote

$$\Lambda^0(g) := \sup_p \{ \langle p, g \rangle - D(p \| p^0) \} \quad (23)$$

Lemma 1.1 (Convex dual of relative entropy). *For each $p^0 \in \mathcal{P}(\mathcal{X}^{M+1})$ and function $g: \mathcal{X}^{M+1} \rightarrow \mathbb{R}$, the (possibly infinite) value of (23) coincides with the log moment generating function:*

$$\Lambda^0(g) = \log \langle p^0, e^g \rangle$$

Moreover, provided $\Lambda^0(g) < \infty$, the supremum in (23) is uniquely attained at $p^ = p^0 \exp(g - \Lambda^0(g))$. That is, the log-likelihood $L^* = \log(p^*/p^0)$ is given by*

$$L^*(\vec{x}) = g(\vec{x}) - \Lambda_0(g).$$

Proof of Prop. 3.1. The optimization problem that defines φ^* admits a convenient decomposition:

$$\begin{aligned} \varphi^*(\beta, h) &= - \sup_p \left\{ \langle p, g_\beta \rangle - D(p \| p^0) \right\} \\ &\quad + \min_\gamma \left\{ \frac{\kappa}{2} \|\gamma\|^2 + \beta^T[\gamma + r] \right\} - \langle p^0, h \rangle \end{aligned} \quad (24)$$

where $g_\beta = \mathcal{Y}_\beta - h$. Lemma 1.1 then gives

$$\varphi^*(\beta, h) = -\Lambda_0(g_\beta) + \min_\gamma \left[\frac{\kappa}{2} \|\gamma\|^2 + \beta^T[\gamma + r] \right] - \langle p^0, h \rangle$$

where $\Lambda_0(g_\beta) = \log \langle p^0, e^{\mathcal{Y}_\beta - h} \rangle$. Minimizing over γ gives $\kappa\gamma^* = -\beta$, and substituting into the above gives (i).

The function to be maximized in part (ii) is denoted

$$\mathcal{G}(h) = -\log \langle p^0, e^{\mathcal{Y}_\beta - h} \rangle - \langle p^0, h \rangle$$

Since h depends only on x_0 , we obtain

$$\mathcal{G}(h) = -\log \langle \nu_0^0, e^g \rangle - \langle \nu_0^0, h \rangle$$

where $g(x_0) = -h(x_0) + \log \left(\sum_{x_1, \dots, x_M} p^0(\vec{x} | x_0) e^{\mathcal{Y}_\beta(x)} \right)$. Recall that $\log \langle \nu_0^0, e^g \rangle = \Lambda_0(g)$ is a log moment-generating function and hence Lemma 1.1 implies the inequality

$$\Lambda_0(g) \geq \Lambda_0(g_0) + \langle \nu_0^{g_0}, g - g_0 \rangle$$

where $g - g_0 = h_0 - h$. Substitution into $\mathcal{G}(h) = -\Lambda_0(g) - \langle \nu_0^0, h \rangle$ reveals another inequality

$$\mathcal{G}(h) \leq -\Lambda_0(g_0) - \langle \nu_0^{g_0}, h_0 - h \rangle - \langle \nu_0^0, h \rangle$$

Adding and subtracting $\langle \nu_0^0, h_0 \rangle$ results in

$$\mathcal{G}(h) \leq \mathcal{G}(h_0) + \langle \nu_0^{g_0} - \nu_0^0, h - h_0 \rangle$$

We see that h_0 is a maximizer when $\nu_0^{g_0} - \nu_0^0 = 0$. We can now characterize the optimizer:

$$g^* = g_0 \quad \text{provided} \quad \nu_0^{g_0} = \nu_0^0 \quad (25)$$

Since $\nu_0^{g^*} = \nu_0^0 \exp(g^* - \Lambda_0(g^*))$, we must have $g^* \equiv \Lambda_0(g^*)$, which is satisfied by any constant function g^* . For simplicity, we take $0 = g^* = \Lambda_0(g^*)$, resulting in (ii):

$$h_\beta^*(x_0) = \log \left(\sum_{x_1, \dots, x_M} p^0(\vec{x} | x_0) e^{\mathcal{Y}_\beta(x)} \right)$$

The proof of (iii) can be obtained from Lemma 3.3: $\partial^2 \varphi^*(\beta) \leq -\kappa^{-1} I$ for every β , which implies that φ^* is strictly concave. \square

B. Proof of Prop. 3.2

Proof. First, we prove (21), then we show each P_k^β is a Markov transition matrix. Thm. 2.1 allows p^β to be expressed

$$p^\beta(\vec{x}) = p^0(\vec{x}) \exp\left(\sum_{k=1}^M \beta_k \mathcal{Y}(x_k) - \Lambda_\beta(x_0)\right), \quad (26)$$

which leads to

$$p^\beta(\vec{x}) = \nu_0^0(x_0) \exp(-\Lambda_\beta(x_0)) \widehat{P}_0(x_0, x_1) \widehat{P}_1(x_1, x_2) \cdots \quad (27)$$

Next, we insert the functions a_1 through a_M , while leaving $p^\beta(x)$ unchanged:

$$\begin{aligned} p^\beta(x) &= \nu_0^0(x_0) \exp(-\Lambda_\beta(x_0)) \widehat{P}_0(x_0, x_1) a_1(x_1) \\ &\quad \frac{1}{a_1(x_1)} \widehat{P}_1(x_1, x_2) a_2(x_2) \frac{1}{a_2(x_2)} \cdots \\ &= \nu_0^0(x_0) \exp(-\Lambda_\beta(x_0)) P_0(x_0, x_1) a_1(x_1) \\ &\quad P_1^\beta(x_1, x_2) P_2^\beta(x_2, x_3) \cdots \end{aligned}$$

Via induction, we see that $a_0(x_0)$ can be calculated via the matrix product:

$$a_0(x_0) = \left[\prod_{k=0}^{M-1} \widehat{P}_k \right] \mathbf{1} \Big|_{x=x_0} \quad (28)$$

where $\mathbf{1}$ is a column vector of ones.

Now consider the term $\exp(\Lambda_\beta(x_0))$. From (12) and repeated use of the Markov property:

$$\exp(\Lambda_\beta(x_0)) = \left[\prod_{k=0}^{M-1} \widehat{P}_k \right] \mathbf{1} \Big|_{x=x_0} \quad (29)$$

Thus, $\exp(-\Lambda_\beta(x_0)) = 1/a_0(x_0)$ and (21) is true.

Now we will show each $P_k^\beta(x, x')$ is a Markov transition matrix. Consider the definition of $P_k^\beta(x, x')$:

$$P_k^\beta(x, x') = \frac{1}{a_k(x)} \widehat{P}_k(x, x') a_{k+1}(x'), \quad (30)$$

Each element of $P_k^\beta(x, x')$ is non-negative because each element of $a_k(x)$, $a_{k+1}(x')$, and $\widehat{P}_k(x, x')$ are non-negative. The rows also sum to one:

$$\sum_{x'} P_k^\beta(x, x') = \frac{1}{a_k(x)} \sum_{x'} \widehat{P}_k(x, x') a_{k+1}(x') = 1 \quad (31)$$

Thus, p^β is Markov for any β . \square

REFERENCES

- [1] S. Meyn, P. Barooah, A. Bušić, Y. Chen, and J. Ehren, "Ancillary service to the grid using intelligent deferrable loads," *IEEE Trans. Automat. Control*, vol. 60, no. 11, pp. 2847–2862, Nov 2015.
- [2] Y. Chen, M. U. Hashmi, J. Mathias, A. Bušić, and S. Meyn, "Distributed control design for balancing the grid using flexible loads," in *Energy Markets and Responsive Grids: Modeling, Control, and Optimization*, S. Meyn, T. Samad, I. Hiskens, and J. Stoustrup, Eds. New York, NY: Springer, 2018, pp. 383–411.
- [3] A. Bušić and S. Meyn, "Distributed randomized control for demand dispatch," in *Proc. of the IEEE Conf. on Dec. and Control*, Dec 2016, pp. 6964–6971.
- [4] J. Mathias, A. Bušić, and S. Meyn, "Demand dispatch with heterogeneous intelligent loads," in *Proc. 50th Annual Hawaii International Conference on System Sciences (HICSS)*.
- [5] N. Cammardella, J. Mathias, M. Kiener, A. Bušić, and S. Meyn, "Balancing California's grid without batteries," in *Proc. of the IEEE Conf. on Dec. and Control*, Dec 2018, pp. 7314–7321.
- [6] H. Hao, B. Sanandaji, K. Poolla, and T. Vincent, "A generalized battery model of a collection of thermostatically controlled loads for providing ancillary service," in *51st Annual Allerton Conference on Communication, Control, and Computing*, Oct 2013, pp. 551–558.
- [7] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent, "Aggregate flexibility of thermostatically controlled loads," *IEEE Trans. on Power Systems*, vol. 30, no. 1, pp. 189–198, Jan 2015.
- [8] H. Hao, T. Middelkoop, P. Barooah, and S. Meyn, "How demand response from commercial buildings will provide the regulation needs of the grid," in *50th Allerton Conference on Communication, Control, and Computing*, 2012, pp. 1908–1913.
- [9] S. Meyn, P. Barooah, A. Bušić, and J. Ehren, "Ancillary service to the grid from deferrable loads: The case for intelligent pool pumps in Florida," in *Proc. of the IEEE Conf. on Dec. and Control*, Dec 2013, pp. 6946–6953.
- [10] R. Malhamé and C.-Y. Chong, "Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system," *IEEE Trans. Automat. Control*, vol. 30, pp. 854–860, 1985.
- [11] R. Malhamé and C.-Y. Chong, "On the statistical properties of a cyclic diffusion process arising in the modeling of thermostat-controlled electric power system loads," *SIAM J. Appl. Math.*, vol. 48, no. 2, pp. 465–480, 1988.
- [12] A. Kizilkale and R. Malhamé, "A class of collective target tracking problems in energy systems: Cooperative versus non-cooperative mean field control solutions," in *Proc. of the IEEE Conf. on Dec. and Control*, 2014, pp. 3493–3498.
- [13] A. Kizilkale and R. Malhamé, "Collective target tracking mean field control for Markovian jump-driven models of electric water heating loads," in *Control of Complex Systems*, K. G. Vamvoudakis and S. Jagannathan, Eds. Butterworth-Heinemann, 2016, pp. 559 – 584.
- [14] M. Almassalkhi, J. Frolik, and P. Hines. Packetized energy management: asynchronous and anonymous coordination of thermostatically controlled loads. In *Proc. of the American Control Conf.*, pages 1431–1437. IEEE, 2017.
- [15] D. Métivier, I. Luchnikov, and M. Chertkov, "Power of ensemble diversity and randomization for energy aggregation," *Scientific Reports*, vol. 9, no. 1, p. 5910, 2019.
- [16] M. Kárný, "Towards fully probabilistic control design," *Automatica*, vol. 32, no. 12, pp. 1719–1722, 1996.
- [17] E. Todorov, "Linearly-solvable Markov decision problems," in *Advances in Neural Information Processing Systems 19*, B. Schölkopf, J. Platt, and T. Hoffman, Eds. MIT Press, 2007, pp. 1369–1376.
- [18] P. Guan, M. Raginsky, and R. Willett, "Online Markov decision processes with Kullback-Leibler control cost," *IEEE Trans. Automat. Control*, vol. 59, no. 6, pp. 1423–1438, June 2014.
- [19] M. Chertkov and V. Y. Chernyak, "Ensemble control of cycling energy loads: Markov Decision Approach," in *IMA volume on the control of energy markets and grids*. Springer, 2017.
- [20] A. Dembo and O. Zeitouni, *Large Deviations Techniques And Applications*, 2nd ed. New York: Springer-Verlag, 1998.
- [21] S. Boyd and L. Vandenberghe, *Convex Optimization*, 1st ed. New York: Cambridge University Press, 2004.
- [22] D. G. Luenberger, *Optimization by vector space methods*. New York: John Wiley & Sons Inc., 1969, reprinted 1997.
- [23] A. Bušić and S. Meyn, "Ordinary Differential Equation Methods for Markov Decision Processes and Application to Kullback–Leibler Control Cost," *SIAM J. Control Optim.*, vol. 56, pp. 343–366, 2018.
- [24] —, "Action-Constrained Markov Decision Processes With Kullback-Leibler Cost," in *Proceedings of the Conference on Computational Learning Theory, and ArXiv e-prints*, 2018.
- [25] —, "Distributed control of thermostatically controlled loads: Kullback-Leibler optimal control in continuous time," in *Proc. of the IEEE Conf. on Dec. and Control*, Dec 2019, p. PP.
- [26] R.-R. Chen and S. P. Meyn, "Value iteration and optimization of multiclass queueing networks," *Queueing Syst. Theory Appl.*, vol. 32, no. 1-3, pp. 65–97, 1999.