Distributed Control of Thermostatically Controlled Loads: Kullback-Leibler Optimal Control in Continuous Time

Ana Bušić, Sean Meyn

Distributed Control of Thermostatically Controlled Loads: Kullback-Leibler Optimal Control in Continuous Time

Ana Bušić[1] and Sean Meyn[2]

***Abstract*— The paper develops distributed control techniques to obtain grid services from flexible loads. The *Individual Perspective Design* (IPD) for local (load level) control is extended to piecewise deterministic and diffusion models for thermostatically controlled load models.**

The IPD design is formulated as an infinite horizon average reward optimal control problem, in which the reward function contains a term that uses relative entropy rate to model deviation from nominal dynamics. In the piecewise deterministic model, the optimal solution is obtained via the solution to an eigenfunction problem, similar to what is obtained in prior work. For a jump diffusion model this simple structure is absent. The structure for the optimal solution is obtained, which suggests an ODE technique for computation that is likely far more efficient than policy- or value-iteration.

I. Introduction

This paper concerns distributed control of a large collection of electric loads, focusing on thermostatically controlled loads (TCLs). This is a topic of significant recent research [1], [5], [6], [20]–[22], [24]; the motivation of each of these authors is to obtain grid level services from the collection of loads, by harnessing their inherent flexibility.

A high level description is based on a single *balancing authority* (BA) that must meet supply-demand balance in the grid in real-time. The BA has access to controllable generation and battery systems for this purpose. In addition, a large collection of loads can receive instructions from the BA to increase/decrease power consumption in such a way that the ensemble appears as a large battery. The creation of *virtual energy storage* from a collection of loads requires either point-to-point control of each load, as in the priority-stack approach [11], or a carefully designed distributed control architecture.

The current paper extends the control architecture introduced in [21], which in its most basic form is composed of two parts:

(i) *Intelligence at the load:* at each time t, the ith load observes its internal state $X^i(t)$ along with a common signal ζ_t broadcast from the BA to all loads of a given type. The load changes its power consumption with some probability that depends on the observations $(X^i(t), \zeta_t)$.

(ii) *Intelligence at the BA:* Given aggregate measurements of the grid and the load population, the signal $\boldsymbol{\zeta}$ is synthesized. This requires a macro model of input-output behavior, where $\boldsymbol{\zeta}$ is the input and the output $\boldsymbol{y}$ is the aggregate power deviation.

A macro-model for a population of loads begins with an understanding of the individual. This leads to a mean-field model, which was originally introduced to the power systems field in [17], [18]. Models for an individual TCL are well understood, and are typically formulated as a first-order linear system with "jumps" due to hysteresis. In the case of a single water heater, a typical model is

$$\frac{d}{dt}\Theta(t) = -\beta[\Theta(t) - \Theta^a(t)] + \gamma M(t) - d(t)\,, \qquad (1)$$

with $d(t)$ is a function of inlet flow and the temperature of the cold water entering the tank, $\Theta(t)$ is the temperature of the water in the tank, $\Theta^a(t)$ is ambient temperature (deg. F.), and β, γ are constants [5], [24].

The process $\boldsymbol{M} = \{M(t)\}$ is the power mode: taking values 0 or 1. A hysteresis control rule is determined by a temperature interval $[\theta_{\min}, \theta_{\max}]$. The power mode changes only when $\Theta(t)$ reaches the boundary of this interval. For example, when $\Theta(t_0) = \theta_{\max}$, then $M(t_0) = 0$ so that the temperature may cool to its lower limit.

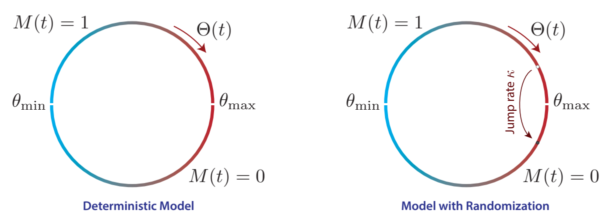

Fig. 1: State space for a water heater under hysteresis control: left hand side shows behavior with deterministic hysteresis control, and right hand side shows behavior under local randomized control

A challenge with TCL models is that the state space is continuous, while the local control designs in much of prior research requires a discrete state space (this is true in particular for [20], [21]). The papers [6], [24] embrace the continuous state space, but their control solutions present difficulties discussed below. Malhamé introduced mean field models for TCLs in continuous time [17]. More recently, he and his co-authors have introduced mean field game techniques for control: [12], [13] contain surveys. The approach is open loop, but the loop can be closed through an MPC implementation.

The main contribution of this paper is to demonstrate that the local control designs introduced in [3], [4], [21] admit

Funding from the ARPA-E NODES award, *Open and Scalable Distributed Energy Resource Networks*, National Science Foundation award EPCN 1609131, and French National Research Agency grant ANR-16-CE05-0008 are gratefully acknowledged.

[1]Inria and DI ENS, École Normale Supérieure, CNRS, PSL Research University, Paris, France

[2]Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, and Inria International Chair

practical extension to continuous-state models. This is true even in the significantly more complex setting in which the nominal dynamics include stochastic disturbances that are outside of direct control (such as variations in ambient temperature, or inlet temperature of water to the TCL).

A. Related Literature

Mean-field models for TCLs have been considered since the 1994 paper by Laurent and Malhamé [15]. The jump-diffusion model considered in Section IV is similar to the model of [15], [16], except that in this prior work the jump rate is assumed independent of temperature.

The most closely related work is contained in the book chapter [6], and in particular a problem formulation proposed in Section 5 of this prior work. The finite time-horizon problem is considered:

$$\begin{aligned}&\min_{\boldsymbol{p}} D_T(\boldsymbol{p}\|\boldsymbol{p}^0)\\ &\text{s.t. } \mathsf{E}_p[\mathcal{U}(X(t))] = r_t\,, \qquad 0\le t\le T\end{aligned} \tag{2}$$

where D_T denotes relative entropy, and $\boldsymbol{r}$ is a reference signal to be tracked, and $\mathcal{U}(x)$ denotes power consumption in state x. Relative entropy is a useful notion of distance between probability measures, and its surrounding geometry leads to tractable algorithms in optimization (as will be seen in the present paper).

While elegant, the approach does present challenges. First, it is an open-loop solution: the authors of [6] suggest a forward-backward (in time) algorithm to compute the optimizer. It requires knowledge of $\boldsymbol{p}^0$, which means knowledge of the initial histogram of loads. The control solution is only physically meaningful for a model in which all randomness is introduced through randomized control. The optimal solution to (2) will distort the dynamics of each load, without respect for the laws of physics: this is made clear in Lemma 4.1 for the diffusion model considered here. The conclusion also holds for models in discrete time. If constraints on the optimization problem are imposed so that physical constraints are respected, then the optimizer can be efficiently computed only in special settings [4], [5]. Finally, the control objective can be viewed as “dead beat” control, which may be infeasible or sensitive to model error.

Another related paper is [24], which treats TCL models in continuous time. This paper is in fact more similar to the prior work [19], [20]. Through a sequence of transformations, the authors obtain a *linear model* for input-output dynamics. In [24] the authors propose dead-beat control, as in the control formulation (2), while [19], [20] propose LQR control. A challenge with these approaches is implementation, since the original model is far from linear. In each of these papers, once the control solution is found for the linear model, this must be transformed to obtain the randomized decision rule for the load. This step requires, for each time t, division by an estimate of the histogram of loads. The robustness of this control approach has not been tested.

The remainder of this paper is organized as follows. Section II contains notation and the model description. The IPD solution is obtained first in Section III for the Piecewise Deterministic Markov Model, and then generalized in Section IV to a diffusion model. Conclusions and directions for future research are contained in Section V.

II. Distributed Control Architecture

The state process for a load at time t is expressed

$$X(t) = (M(t), \Theta(t))\,, \qquad t\ge 0 \tag{3}$$

which evolves on the state space $\mathsf{X} = \{0,1\}\times\mathbb{R}$. Further constraints are imposed in the piece-wise deterministic model considered in Section III, so that $\boldsymbol{X}$ evolves on the compact set $\mathsf{X} = \{0,1\}\times[\theta_{\min},\theta_{\max}]$. In this case, it is convenient to identify the two bounded intervals with a circle, as illustrated in Fig. 1.

The mean-field model over a time horizon $[0,T]$ is defined by a probability measure $\boldsymbol{p}$ on state-trajectories $\boldsymbol{X} = \{X(t) : 0\le t\le T\}$. At each time t, the marginal p_t is intended to approximate the histogram of the finite population: for a continuous function $f\colon \mathsf{X}\to\mathbb{R}$,

$$\frac{1}{N}\sum_{i=1}^{N} f(X^i(t)) \approx \langle p_t, f\rangle := \int_{\mathsf{X}} f(x)p_t(x)\,dx$$

In prior work [5], [12], [13], [17], [18], assumptions are imposed so that this approximation becomes exact in the limit as $N\to\infty$.

The decentralized control approach proposed in the prior work [3], [5], [21] is based on the construction of a controlled Markov model with input denoted $\boldsymbol{\zeta}$. This may be regarded as a randomized decision rule at each load, designed to respect quality of service (QoS) constraints, and also ensure that the input-output behavior from $\boldsymbol{\zeta}$ to $\boldsymbol{y}$ has desirable properties from the viewpoint of the BA.

The design of the parametrized family of randomized decision rules is just the first step in a distributed control architecture. The balancing authority broadcasts the common signal $\boldsymbol{\zeta} = \{\zeta_t : t\ge 0\}$ to all loads, based on measurements of aggregate power consumption $\boldsymbol{y} = \{y_t : t\ge 0\}$.

A. Local control

Consider an aggregator in a region of at least one hundred homes, with electric loads such as water heaters, air-conditioning, and refrigerators. For each type of load, the aggregator wishes to ramp up and down power consumption to meet the needs of the grid, subject to strict bounds on QoS delivered by each load to each household. For example, in the case of a water heater, the temperature must remain within pre-determined bounds. Additional on/off cycling is permitted, but excess cycling may also be subject to strict bounds.

Local control refers to the randomized decision rule at an individual load. This is specified by modification of the statistics of the power mode process $\boldsymbol{M}$, and analysis is based on the associated differential generator.

The main contribution of this paper is local control design for models in continuous time and continuous state. This avoids an approximation based on quantization of the state space, and theory gives greater insight on the structure of the solutions. The theory is developed for a jump diffusion model for which a PDE approach can be used to generate the family of differential generators. Their dependence on ζ is only through the modified jump-rate function. An understanding of this transformation is a central technical component of the paper.

B. Markov models

The notation $x = (m, \theta)$ is used for any state $x \in \mathsf{X}$, where $m \in \{0, 1\}$ and $\theta \in \mathbb{R}$. For $m \in \{0, 1\}$, denote by $\overline{m} \in \{0, 1\}$ the power mode that is alternate to m; that is, $\overline{m} = 1 - m$. For any function $g\colon \mathsf{X} \to \mathbb{R}$, denote

$$\bar{g}(m, \theta) = g(\overline{m}, \theta) , \quad x = (m, \theta) \in \mathsf{X} . \tag{4}$$

A Markovian model for a $\boldsymbol{X}$ is described by a differential generator $\mathcal{D}$. For all functions $g\colon \mathsf{X} \to \mathbb{R}$ in some domain,

$$\mathcal{D}g\,(x) = \lim_{t\downarrow 0} \frac{1}{t} \mathsf{E}[g(X(t)) - g(X(0)) \mid X(0) = x] \tag{5}$$

The most general form considered is expressed as follows:

$$\mathcal{D}g = \mathcal{Q}_0 g + \kappa[\bar{g} - g] , \tag{6}$$

where $\kappa\colon \mathsf{X} \to \mathbb{R}_+$ is the jump-rate function (JRF). The nominal model with generator $\mathcal{D}_0$ has JRF denoted κ_0. In control solutions it is permissible to modify κ_0, but not $\mathcal{Q}_0$.

For simplicity it is assumed in this paper that $\mathcal{Q}_0$ is a first or second-order differential operator,

$$\mathcal{D}g = \rho g' + \tfrac{1}{2}\sigma_B^2 g'' + \kappa[\bar{g} - g] . \tag{7}$$

The drift function ρ is described in Section III. The instantaneous variance σ_B^2 is part of the diffusion model described in Section IV.

It is assumed that there is a unique invariant density ϱ_0 for the nominal model. The steady-state mean of a function $g\colon \mathsf{X} \to \mathbb{R}$ is denoted

$$\langle \varrho_0, g \rangle := \int g(x) \varrho_0(x)\, dx$$

Invariance implies that $\langle \varrho_0, \mathcal{D}_0 g \rangle = 0$ for $g \in C_0^2$.

The function $\mathcal{U}\colon \mathsf{X} \to \mathbb{R}_+$ denotes power mode: $\mathcal{U}(x) = m$ for $x = (m, \theta) \in \mathsf{X}$. Poisson's equation for the nominal model is

$$\mathcal{D}_0 H = -\widetilde{\mathcal{U}} \tag{8}$$

in which H is the solution, and $\widetilde{\mathcal{U}} = \mathcal{U} - \bar{u}_0$ is normalized so it has zero steady-state mean: $\bar{u}_0 = \langle \varrho_0, \mathcal{U} \rangle$.

Distributed control solutions will be based on a family of models, parameterized by a scalar $\zeta \in \mathbb{R}$, with similar notational conventions for the generator and other objects of interest: $\check{\mathcal{D}}_\zeta$, $\check{\varrho}_\zeta$, and $\bar{u}_\zeta := \langle \check{\varrho}_\zeta, \mathcal{U} \rangle$. A typical design results in a generator of the form:

$$\check{\mathcal{D}} := I_v^{-1}[I_G + \mathcal{D}_0] I_v , \tag{9}$$

in which G, v are C^2 functions on X, and I_G denotes the multiplication operator. The constraint $\check{\mathcal{D}}\mathbb{1} = 0$ results in $G = -v^{-1}\mathcal{D}_0 v$.

C. Individual perspective design

The Individual Perspective Design (IPD) for Markov models in discrete time was developed in [4], [5], [21]. Its continuous time version is defined by the solution to the convex program,

$$\arg\max_{\boldsymbol{p}} \Big\{ \zeta \mathsf{E}_p \Big[\int_0^T \mathcal{U}(X(t))\, dt \Big] - D_T(\boldsymbol{p} \| \boldsymbol{p}^0) \Big\} \tag{10}$$

where the expectation is with respect to the probability law $\boldsymbol{p}$, and D_T again denotes relative entropy. Attention here is devoted to the infinite horizon case, based on the *relative entropy rate*:

$$\mathcal{R}(\mathcal{D} \| \mathcal{D}_0) := \limsup_{T\to\infty} \frac{1}{T} D_T(\boldsymbol{p} \| \boldsymbol{p}^0) . \tag{11}$$

Following the definition of the IPD solution of [2], [3], [21], we denote

$$\eta_\zeta^* := \max_{\varrho, \mathcal{D}} \Big\{ \zeta \langle \varrho, \mathcal{U} \rangle - \mathcal{R}(\mathcal{D} \| \mathcal{D}_0) \Big\} \tag{12}$$

subject to the invariance constraint $\varrho\mathcal{D} = 0$. The optimal controlled generators are denoted $\{\check{\mathcal{D}}_\zeta : \zeta \in \mathbb{R}\}$.

An alternative interpretation of the IPD is through a constrained optimization problem. The proof of Prop. 2.1 is via a Lagrangian relaxation of the constraint $\langle \varrho, \mathcal{U} \rangle = r$.

Proposition 2.1: For any $\zeta \in \mathbb{R}$, let $\check{\mathcal{D}}_\zeta, \check{\varrho}_\zeta$ denote a solution to (12) and let r denote the mean

$$r = \bar{u}_\zeta := \langle \check{\varrho}_\zeta, \mathcal{U} \rangle \tag{13}$$

Then $\check{\mathcal{D}}_\zeta, \check{\varrho}_\zeta$ also optimize the convex program

$$\begin{aligned} &\min_{\varrho, \mathcal{D}} \quad \mathcal{R}(\mathcal{D} \| \mathcal{D}_0) \\ &\text{subject to} \quad \langle \varrho, \mathcal{U} \rangle = r , \quad \varrho\mathcal{D} = 0 \end{aligned} \tag{14}$$

That is, the solutions to (14) and (12) coincide. $\square$

For the piecewise deterministic model considered in Section III, $\check{\mathcal{D}}_\zeta$ is obtained as the solution to an eigenvector problem, exactly as in the discrete-time case [5], [21]. It is shown in Section IV that this approach fails for the diffusion model, but the family of optimizers $\{\check{\mathcal{D}}_\zeta\}$ can be computed through a reformulation of the optimal control problem.

III. Piecewise Deterministic Markov Model

A Markovian model is considered here, in which the only randomness in the system is introduced through the power mode process $\boldsymbol{M}$. It is a piecewise-deterministic Markov process (PDMP) [7], [8]; in Section IV we turn to a jump-diffusion process.

Each of the models considered in this paper are variants of the standard TCL model (1). For the nominal model it is necessary to introduce randomness in the times at which the power mode $\boldsymbol{M}$ changes.

The state space for the PDMP model remains equal to the union of two half-circles. However, with the introduction of a randomized mode process, the state process will jump between the upper and lower halves, as illustrated at right in Fig. 1, rather than flow continuously in the clockwise direction. The randomness will be designed with care, so that cycling is increased by a small amount (say, 5%).

A brief survey of the rate function for a Markov process can be found in Section 2.2 of [26], with more details in [25]. Girsanov's transformation plays a central role in this paper as well as in [3], [4], [6], [21]. Section 6.2 of [25] contains useful material in this continuous time setting; see also Theorem 1.31 of [23] (Girsanov Theorem I for Itô-Lévy Processes).

A. Model construction

A time-invariant approximation for (1) is described as a nonlinear ordinary differential equation:

$$\frac{d}{dt}\Theta(t) = \rho(X(t)), \qquad t \ge 0. \tag{15}$$

In the PDMP it is assumed that the temperature process evolves as (15) in-between jumps, with $\theta_{\min} \le \Theta(t) \le \theta_{\max}$ for all t, so that $\mathsf{X} = \{0,1\} \times [\theta_{\min}, \theta_{\max}]$.

The detailed structure of the function ρ is not important, except for the following controllability property: for some $\varepsilon_\rho > 0$,

$$\rho(1,\theta) \ge \varepsilon_\rho \ \text{ if } \theta < \theta_{\max}, \quad \rho(0,\theta) \le -\varepsilon_\rho \ \text{ if } \theta > \theta_{\min}. \tag{16}$$

It is also assumed that ρ is continuous and bounded on X.

The statistics of jump-times are specified by a jump-rate function (JRF) $\kappa\colon \mathsf{X} \to \mathbb{R}_+$. For each $t \ge 0$, the following approximation holds for small $\delta > 0$:

$$\mathsf{P}\{M(t+\delta) \ne M(t) \mid \mathcal{F}_t\} = \delta\kappa(X(t)) + o(\delta) \tag{17}$$

with filtration generated by the state process:

$$\mathcal{F}_t := \sigma\{X(s) : s \le t\}, \quad t \ge 0 \tag{18}$$

Any JRF κ is assumed to be non-negative and continuous in the interior of X. In addition, the following assumptions will be assumed throughout:

$$\lim_{\theta\uparrow\theta_{\max}} \kappa(1,\theta) = \infty, \quad \lim_{\theta\downarrow\theta_{\min}} \kappa(0,\theta) = \infty, \tag{19}$$

$$\int_{\theta_{\min}}^{\theta_{\max}} \kappa(1,\theta) = \int_{\theta_{\min}}^{\theta_{\max}} \kappa(0,\theta) = \infty. \tag{20}$$

The resulting jump process never hits a temperature boundary: $\theta_{\min} < \Theta(t) < \theta_{\max}$ for all $t > 0$.

To avoid frequent cycling it is reasonable to impose the constraint $\kappa(1,\theta) = 0$ in a neighborhood of $\theta_{\min}$, and $\kappa(0,\theta) = 0$ in a neighborhood of $\theta_{\max}$.

The JRF for the nominal model is denoted κ_0.

B. Realizations

To describe a realization of the Markov process with JRF κ, it suffices to describe the statistics of the first jump time:

$$\tau_1 = \min\{t > 0 : M(t) \ne M(0)\}. \tag{21}$$

For each $x = (m,\theta) \in \mathsf{X}$, let $\Psi(t;x)$ denote the solution to the ordinary differential equation:

$$\Psi(t;x) = \xi(t), \ \frac{d}{dt}\xi(t) = \rho(\xi(t), m), \quad t \ge 0, \ \xi(0) = \theta, \tag{22}$$

and denote $\bar\tau_1 = \min\{t \ge 0 : \Psi(t;x) \in \partial\mathsf{X}\}$, where the boundary is the finite set $\partial\mathsf{X} = \{0,1\} \times \{\theta_{\min}, \theta_{\max}\}$.

Let R^* denote a unit-mean exponentially distributed random variable that is independent of $X(0) = x_0$, and denote

$$R(t) = \int_0^t \kappa(\Psi(s;x_0))\, ds, \qquad 0 \le t < \bar\tau_1(x_0). \tag{23}$$

The following definition is then consistent with (17):

$$\tau_1 = \min\{t > 0 : R(t) = R^*\}.$$

The bound $\tau_1 < \bar\tau_1$ holds with probability one since the JRF satisfies (19), (20).

This provides a realization of the first jump time. For the initial condition $x_0 = (m_0, \theta_0)$ and all $t \in [0, \tau_1)$, we then define $M(t) = m_0$, $\Theta(t) = \Psi(t;x_0)$, and hence $\frac{d}{dt}\Theta(t) = \rho(X(t))$. This construction can be repeated to define the process for all $t \ge 0$.

C. Local control solutions

Based on the definition (5), it appears that the function $\mathcal{D}g$ can be expressed in the form (6). This is true provided g is smooth, and the additional boundary conditions hold:

$$\begin{aligned} g(0,\theta_{\min}) &:= \lim_{\theta\downarrow\theta_{\min}} g(0,\theta) = g(1,\theta_{\min}) \\ g(1,\theta_{\max}) &:= \lim_{\theta\uparrow\theta_{\min}} g(1,\theta) = g(0,\theta_{\max}) \end{aligned} \tag{24}$$

The boundary conditions are explained in the Appendix. Under these conditions we can justify the expected representation:

$$\mathcal{D}g = \rho g' + \kappa[\bar g - g], \qquad g \in C^1. \tag{25}$$

1) Myopic design: In analogy with the definition in [5], the myopic design is defined here as the family of Markov processes with differential generator of the form (9) in which $v = v_\zeta = e^{\zeta\mathcal{U}}$ for $\zeta \in \mathbb{R}$. It is indeed of the form (6), with modified JRF:

Proposition 3.1: The myopic design has generator defined for C^1 functions g by $\check{\mathcal{D}}_\zeta g = \rho g' + \kappa_\zeta[\bar g - g]$, with

$$\kappa_\zeta = \exp(\zeta[\bar{\mathcal{U}} - \mathcal{U}])\kappa_0 \tag{26}$$

□

The same conclusion holds for the diffusion model; see Prop. 4.2. So, we postpone the proof to the next section.

2) IPD design: The convex optimization problem (12) defines an optimal JRF denoted $\check\kappa_\zeta$, which then defines a generator $\check{\mathcal{D}}_\zeta$ of the form (25) in which ρ is unchanged.

As in the discrete-time setting [4], [5], [21], the IPD solution is based on the associated eigenfunction problem:

$$[\zeta I_{\mathcal{U}} + \mathcal{D}_0]v_\zeta = \Lambda_\zeta v_\zeta, \quad \zeta \in \mathbb{R}, \tag{27}$$

which for this model becomes

$$\zeta\mathcal{U}v_\zeta + \rho v_\zeta' + \kappa_0[\bar v_\zeta - v_\zeta] = \Lambda_\zeta v_\zeta \tag{28}$$

For fixed ζ, if Λ_ζ is given, then v_ζ is obtained as the solution to two coupled ODEs.

The solution is unique only up to a multiplicative constant: fix a state $\theta^\circ \in \mathsf{X}$, and choose the solution to satisfy $v_\zeta(\theta^\circ) = 1$ for all ζ.

Theorem 3.2: Suppose that the eigenfunction equation (27) admits a family of solutions $\{v_\zeta\}$ that are continuously differentiable in ζ. Then, the solution to the optimal-reward optimization problem has the following structure for the PDMP model. For each $\zeta \in \mathbb{R}$,

(i) The optimal JRF is

$$\check{\kappa}_\zeta(x) = \kappa_0(x)\frac{v_\zeta(\bar{x})}{v_\zeta(x)}\,, \quad x \in \mathsf{X}\,. \tag{29}$$

(ii) The function $H^*_\zeta := \frac{d}{d\zeta}\log(v^*_\zeta)$ solves Poisson's equation:

$$\check{\mathcal{D}}_\zeta H^*_\zeta = -\widetilde{\mathcal{U}}_\zeta\,, \qquad \widetilde{\mathcal{U}}_\zeta = \mathcal{U} - \bar{u}_\zeta \tag{30}$$

(iii) $\eta^*_\zeta = \Lambda_\zeta$. $\square$

The proof is contained in the Appendix.

Smoothness of $\{v_\zeta\}$ in ζ is established in [4] for a discrete-time/discrete-space model. The proof, based on the implicit function theorem, likely carries over to the present setting.

IV. Diffusion Model

Consider the following extension of the PDMP model. Within an interval of continuity, the temperature evolves as the SDE

$$d\Theta(t) = \rho(X(t))\,dt + dB(t) \tag{31}$$

in which ρ is as in (15), and $\boldsymbol{B}$ is Brownian motion, with instantaneous variance denoted $\sigma^2_B \ge 0$. It is assumed that $X(t)$ is independent of $B(t+T) - B(t)$ for each $t, T \ge 0$.

Note that (31) includes the PDMP model as a special case, since we do not rule out $\sigma^2_B = 0$.

The previous assumptions on ρ are maintained: this function is bounded, continuous, and satisfies (16). The state space however is no longer bounded. It is assumed of the form

$$\mathsf{X} = \{0\} \times [\theta_{\min}, \infty) \cup \{1\} \times (-\infty, \theta_{\max}] \tag{32}$$

It is assumed that $M(t) = 0$ when $\Theta(t) > \theta_{\max}$, and $M(t) = 1$ when $\Theta(t) < \theta_{\min}$.

The differential generator is modified to include a second order differential term:

$$\mathcal{D}g = \tfrac{1}{2}\sigma^2_B g'' + \rho g' + \kappa[\bar{g} - g]\,, \quad g \in C^2\,.$$

In the notation introduced in (6), we have $\mathcal{Q}_0 g = \frac{1}{2}\sigma^2_B g'' + \rho g'$ (see [23, Thm. 1.22]). It will be clear that the assumptions on $\mathcal{Q}_0$ are far stronger than necessary. We could for example allow a variance term that depends on the state.

For any JRF under consideration, it is assumed that $\kappa(0,\theta) = 0$ when $\theta > \theta_{\max}$, and $\kappa(1,\theta) = 0$ when $\theta < \theta_{\min}$. It is always assumed that g satisfies (24).

The realization of the Markov process can be described as in Section III-B with only one change. The mapping Ψ is now stochastic: a standard Brownian motion $\boldsymbol{B}$ is given, and for each $x \in \mathsf{X}$, denote the solution to the SDE without jumps by

$$\Psi(t;x) = \xi(t)\,, \quad d\xi(t) = \rho(\xi(t), m)\,dt + dB(t) \tag{33}$$

The first jump time τ_1 is again defined by (21). It does not admit a deterministic upper bound when $\sigma^2_B > 0$.

A. *Preserving dynamic constraints*

Here it is shown that a generator of the form (9) cannot be optimal if the function v depends upon θ. Consider a general transformation of the generator of this form (9), where $v\colon \mathsf{X} \to (0,\infty)$ is a C^2 function. The proof of the following representation follows from the definitions:

Lemma 4.1: The identity $\check{\mathcal{D}}g = \check{\mathcal{Q}}_0 g + \check{\kappa}[\bar{g} - g]$ holds for any C^2 function $g\colon \mathsf{X} \to \mathbb{R}$, where

$$\check{\mathcal{Q}}_0 g = \tfrac{1}{2}\sigma^2_B g'' + \check{\rho} g'\,, \qquad \check{\kappa} = \kappa_0 \bar{v}/v$$
$$\check{\rho}(x) = \rho(x) + \sigma^2_B \tfrac{d}{d\theta}\log(v(x))\,, \qquad x = (m,\theta) \in \mathsf{X}\,.$$

$\square$

If v depends upon θ, then the transformed drift function in the definition of $\check{\mathcal{Q}}_0$ implies altered dynamics:

$$d\Theta(t) = \check{\rho}(X(t))\,dt + dB(t)$$

The solution cannot be applied in practice: *nothing in our modeling assumptions permits us to change ρ.*

B. *Local control*

Given the preliminaries provided in the previous subsection, it is natural to begin with the myopic design.

1) Myopic design: The definition of the myopic design remains the same: generators are the form (9), in which $v = v_\zeta = e^{\zeta\mathcal{U}}$ for $\zeta \in \mathbb{R}$. The proof of Prop. 4.2 follows directly from Prop. 4.1, since v depends only on m.

Proposition 4.2: The myopic design has generator defined for C^1 functions g by

$$\check{\mathcal{D}}_\zeta g = \mathcal{Q}_0 + \kappa_\zeta[\bar{g} - g] \quad \text{where} \quad \kappa_\zeta = e^{\zeta(\bar{\mathcal{U}} - \mathcal{U})}\kappa_0 \tag{34}$$

$\square$

2) IPD solution: Theorem 4.3 provides a representation for the relative entropy rate $\mathcal{R}(\mathcal{D}\|\mathcal{D}_0)$ defined as the limit (11), in which $\boldsymbol{p}^0$ is the distribution for the nominal model, and $\boldsymbol{p}$ is the distribution for the Markov model with differential generator

$$\mathcal{D}g = \mathcal{Q}_0 g + \kappa[\bar{g} - g]\,, \quad g \in C^2\,. \tag{35}$$

Theorem 4.3: the relative entropy rate can be expressed

$$\mathcal{R}(\mathcal{D}\|\mathcal{D}_0) = \int \mathcal{R}_x(\kappa\|\kappa_0)\,\varrho(x)dx \tag{36}$$

where ϱ is the invariant density for the Markov process with generator (35), and

$$\mathcal{R}_x(\kappa\|\kappa_0) = \kappa(x)\log(\kappa(x)/\kappa_0(x)) + \kappa_0(x) - \kappa(x) \tag{37}$$

$\square$

It is a useful fact that $\mathcal{R}(\mathcal{D}\|\mathcal{D}_0)$ depends on σ_B^2 only through the invariant density, and in particular that $\mathcal{R}_x(\kappa\|\kappa_0)$ does not depend on this variance parameter.

The IPD optimization problem (14) is thus the solution to an average-reward optimal control problem. This is true even in the degenerate case $\sigma_B^2 = 0$. The average reward optimality equation (AROE) is

$$\eta_\zeta^* = \max_\kappa \bigl[\zeta\mathcal{U}(x) - \mathcal{R}_x(\kappa\|\kappa_0) + \mathcal{D}h_\zeta^*(x)\bigr]\,, \quad x\in\mathsf{X}\,, \tag{38}$$

where the generator depends on κ via (35). The function $h_\zeta^*\colon \mathsf{X}\to\mathbb{R}$ is known as the *relative value function*. It is not unique, so we normalize as follows: fix a state $\theta^\circ\in\mathsf{X}$, and choose the solution to satisfy $h_\zeta^*(\theta^\circ)=0$ (for each ζ).

Theorem 4.4: Suppose that the AROE admits a family of solutions $\{h_\zeta\}$ that are continuously differentiable in ζ. Then the generator for the optimal solution has the following defining properties:

(i) The maximizer in (38) defines the optimal JRF:

$$\check\kappa_\zeta = \kappa_0 \exp(\bar h_\zeta^* - h_\zeta^*)\,. \tag{39}$$

(ii) The function $H_\zeta^* := \frac{d}{d\zeta}h_\zeta^*$ solves Poisson's equation:

$$\check{\mathcal{D}}_\zeta H_\zeta^* = -\widetilde{\mathcal{U}}_\zeta\,, \qquad \widetilde{\mathcal{U}}_\zeta = \mathcal{U} - \bar u_\zeta \tag{40}$$

These results remain valid in the degenerate case $\sigma_B^2 = 0$. The relative value function in this case is the logarithm of the eigenfunction:

$$h_\zeta^* = \log(v_\zeta) \tag{41}$$

Proof: The formula for $\check\kappa$ is obtained on computing the maximum in (38):

$$\begin{aligned}\check\kappa &= \arg\max_\kappa \bigl[\zeta\mathcal{U}(x) - \mathcal{R}_x(\kappa\|\kappa_0) + \mathcal{D}h^*(x)\bigr]\\ &= \arg\max_\kappa \Bigl[-\kappa(x)\log(\kappa(x)/\kappa_0(x))\\ &\qquad + [\kappa(x)-\kappa_0(x)] + \kappa(x)[h^*(\bar x) - h^*(x)]\Bigr]\end{aligned} \tag{42}$$

where we have substituted the formula (6) for $\mathcal{D}$, and removed terms that do not depend on κ. Applying the first-order conditions for optimality gives (39) and establishes (i).

The identity (41) follows from (i) and Theorem 3.2 (i).

To establish (ii) we simply differentiate each side of (38), in the form

$$\Bigl[\zeta\mathcal{U}(x) - \mathcal{R}_x(\kappa\|\kappa_0) + \mathcal{D}_\kappa h_\zeta^*(x)\Bigr]\Big|_{\kappa=\check\kappa_\zeta} = \eta_\zeta^*\,, \quad x\in\mathsf{X}\,.$$

Observe that (i) implies that $\check\kappa_\zeta$ is C^1 since this is assumed for h_ζ^*. The equation above implies that η_ζ^* is also C^1. This justifies differentiation of each side of this identity. This calculus step is simplified due to the fact that sensitivity with respect to κ is zero:

$$\begin{aligned}&\tfrac{d}{d\zeta}\Bigl[\zeta\mathcal{U}(x) - \mathcal{R}_x(\kappa\|\kappa_0) + \mathcal{D}_\kappa h_\zeta^*(x)\Bigr]\Big|_{\kappa=\check\kappa_\zeta}\\ &\qquad = \mathcal{U}(x) - \mathcal{R}_x(\check\kappa_\zeta\|\kappa_0) + \mathcal{D}_{\check\kappa_\zeta}\tfrac{d}{d\zeta}h_\zeta^*(x)\end{aligned}$$

Substituting $\mathcal{D}_{\check\kappa_\zeta} = \check{\mathcal{D}}_\zeta$ and $\frac{d}{d\zeta}h_\zeta^* = H_\zeta^*$ completes the proof. ∎

V. Conclusions

The treatment of distributed control in continuous time is motivated by the well developed theory of linear TCL models. It is great news that the optimal control solution (the IPD design) has a straightforward extension to continuous time/state models, and that the form of the solution is largely invariant to the form of the model, as seen in Theorem 4.4. Moreover, the solution is in some sense simpler to implement than the solution obtained in the discrete time setting: at the start of a power mode transition, an exponential random variable is drawn. The next power mode transition is made at the time τ_1 given in (21).

Remaining open questions mainly concern implementation. In particular, can the local algorithms be adapted to take into account water usage and other non-stationary disturbances? In the case of water heaters, perhaps these shocks to the system are so rare that they do not impact significantly the behavior of the aggregate. These and many related questions are topics of current research.

References

[1] 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.

[2] A. Bušić and S. Meyn. Passive dynamics in mean field control. In *Proc. of the IEEE Conf. on Dec. and Control*, pages 2716–2721, Dec 2014.

[3] A. Bušić and S. Meyn. Distributed randomized control for demand dispatch. In *Proc. of the IEEE Conf. on Dec. and Control*, pages 6964–6971, Dec 2016.

[4] A. Bušić and S. Meyn. Ordinary Differential Equation Methods for Markov Decision Processes and Application to Kullback–Leibler Control Cost. *SIAM J. Control Optim.*, 56(1):343–366, 2018.

[5] Y. Chen, M. U. Hashmi, J. Mathias, A. Bušić, and S. Meyn. Distributed control design for balancing the grid using flexible loads. In S. Meyn, T. Samad, I. Hiskens, and J. Stoustrup, editors, *Energy Markets and Responsive Grids: Modeling, Control, and Optimization*, pages 383–411. Springer, New York, NY, 2018.

[6] 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 (and ArXiv e-prints)*. Springer, 2017.

[7] M. Davis. Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *J. R. Statist. Soc. B.*, 46:353–388, 1984.

[8] M. Davis. *Markov Models and Optimization*. Chapman and Hall, London, 1993.

[9] D. Down, S. P. Meyn, and R. L. Tweedie. Exponential and uniform ergodicity of Markov processes. *Ann. Probab.*, 23:1671–1691, 1995.

[10] S. N. Ethier and T. G. Kurtz. *Markov Processes : Characterization and Convergence*. John Wiley & Sons, New York, 1986.

[11] 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*, pages 551–558, Oct 2013.

[12] A. Kizilkale and R. Malhame. 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*, pages 3493–3498, 2014.

[13] A. Kizilkale and R. Malhamé. Collective target tracking mean field control for Markovian jump-driven models of electric water heating loads. In K. G. Vamvoudakis and S. Jagannathan, editors, *Control of Complex Systems*, pages 559 – 584. Butterworth-Heinemann, 2016.

[14] I. Kontoyiannis and S. P. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.*, 13:304–362, 2003.

[15] J. C. Laurent and R. P. Malhame. A physically-based computer model of aggregate electric water heating loads. *IEEE Transactions on Power Systems*, 9(3):1209–1217, Aug 1994.

[16] R. Malhamé. A jump-driven Markovian electric load model. *Adv. Appl. Probab.*, 22(3):pp. 564–586, 1990.
[17] 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*, 30(9):854 – 860, Sep 1985.
[18] 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.*, 48:465–480, 1988.
[19] J. Mathieu. *Modeling, Analysis, and Control of Demand Response Resources*. PhD thesis, University of California at Berkeley, 2012.
[20] J. Mathieu, S. Koch, and D. Callaway. State estimation and control of electric loads to manage real-time energy imbalance. *IEEE Trans. Power Systems*, 28(1):430–440, 2013.
[21] S. Meyn, P. Barooah, A. Bušić, Y. Chen, and J. Ehren. Ancillary service to the grid using intelligent deferrable loads. *IEEE Trans. Automat. Control*, 60(11):2847–2862, Nov 2015.
[22] 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*, pages 6946–6953, Dec 2013.
[23] B. K. Øksendal and A. Sulem. *Applied stochastic control of jump diffusions*, volume 498. Springer, 2005.
[24] S. H. Tindemans, V. Trovato, and G. Strbac. Decentralized control of thermostatic loads for flexible demand response. *IEEE Transactions on Control Systems Technology*, 23(5):1685–1700, Sept 2015.
[25] L. Wu. Uniformly integrable operators and large deviations for Markov processes. *Journal of Functional Analysis*, 172(2):301 – 376, 2000.
[26] L. Wu and N. Yao. Large deviation principles for Markov processes via Φ-Sobolev inequalities. *Elec. Commun. Probab.*, 13:10–23, 2008.

APPENDIX

A. Generators

Suppose that $g \in C^2$ and the boundary condition (24) holds. Then, with $f = \mathcal{D}_0 g$, the function below is a *local martingale*:

$$\mathcal{M}(t) := g(X(t)) - g(X(0)) - \int_0^t f(X(s))\, ds \tag{43}$$

If f and g are bounded, then this is a martingale [10].

It is not difficult to see now that the boundary condition (24) is required. To establish the first limit, take an initial condition of the form $x = (0, \theta)$. Suppose that $f = \mathcal{D}_0 g$, with f and g bounded and continuous. Then the following representation follows from the martingale property:

$$\mathsf{E}_x[g(1, \Theta_{\tau_1})] = \mathsf{E}_x[g(X_{\tau_1})] = g(x) + \mathsf{E}_x\Big[\int_0^{\tau_1} f(X_t)\, dt\Big]$$

Consider a sequence of initial conditions of the form $x^k = (0, \theta^k)$ with $\theta^k \downarrow \theta_{\min}$ as $k \to \infty$. Then by (19), (20) we must have $\mathsf{E}_{x^k}[\tau_1] \to 0$, and hence from the above

$$\begin{aligned} g(0, \theta_{\min}) &= \lim_{k\to\infty} g(0, \theta^k) \\ &= \lim_{k\to\infty} \mathsf{E}_{x^k}[g(1, \Theta_{\tau_1})] = g(1, \theta_{\min}) \end{aligned}$$

where the first equation is by continuity of g, the second follows because the expected integral vanishes, and the final limit again uses continuity of g.

B. Ergodic Theorems

All of the models considered in this paper are ergodic in a strong sense. Note that Prop. 1.1 does not exclude $\sigma_B^2 = 0$, so the conclusions hold for the piecewise deterministic model.

Proposition 1.1: For the jump-diffusion with rate function κ and $\sigma_B^2 \geq 0$, there is a C^2 function $V \colon \mathsf{X} \to [1, \infty)$ satisfying the boundary conditions (24), a compact set S, and $\gamma > 0$ such that

$$\mathcal{D}V \leq -\gamma V + b\mathbb{I}_S \tag{44}$$

Moreover, compact sets are *small*, so that the Markov process is V-uniformly ergodic [9].

Proof: (sketch) The proof of the small set condition is not complex, but beyond the scope of this paper because it requires lengthy background from [9].

Turning to the drift condition (44), consider first $V_0(x) = k\log(\exp(\varepsilon_v\theta) + \exp(-\varepsilon_v\theta))$, $x = (m, \theta) \in \mathsf{X}$. For each $\varepsilon_v > 0$ there is $k < \infty$, such that for a compact set S_0 and a constant b_0, a version of Foster's criterion holds: $\mathcal{D}V_0 \leq -1 + b_0\mathbb{I}_{S_0}$. For fixed parameters satisfying this inequality, we take $V = e^{\varepsilon V_0}$, with $\varepsilon > 0$ chosen sufficiently small. ■

A mean ergodic theorem holds for measurable functions L_∞^V, defined as the set of measurable functions $f \colon \mathsf{X} \to \mathbb{R}$ satisfying

$$\|f\|_V := \sup_{x\in\mathsf{X}} \frac{|f(x)|}{V(x)} < \infty\,.$$

For $f \in L_\infty^V$ we have, for each initial condition,

$$\lim_{t\to\infty} \mathsf{E}[f(X(t))] = \langle \varrho, f\rangle$$

where the rate of convergence is exponential [9].

The Law of Large Numbers holds under weaker conditions:

Proposition 1.2: Let ϱ denote the unique invariant measure for the jump-diffusion with rate function κ, and let F, G denote two functions satisfying

$$\langle \varrho, G_-\rangle < \infty\,, \quad \langle \varrho, \kappa F_-\rangle < \infty$$

where $G_- = \max(-G, 0)$, $F_- = \max(-F, 0)$. Then, the following limits hold for each initial condition:

$$\lim_{T\to\infty} \frac{1}{T}\mathsf{E}\Big[\sum_{i:t_i\leq T} F(X(t_i^-))\Big] = \langle \varrho, \kappa F\rangle$$

$$\lim_{T\to\infty} \frac{1}{T}\mathsf{E}\Big[\int_0^T G(X(s))\, ds\Big] = \langle \varrho, G\rangle$$

where $\{t_i : i \geq 1\}$ are the jump times for $\boldsymbol{X}$. □

The notation $X(t_i^-)$ means the usual limit from the left; since it is only the power mode $\boldsymbol{M}$ that jumps, we have

$$X(t_i^-) = \overline{X}(t_i)\,, \quad i \geq 1.$$

C. Infinite-horizon control problem

The theory in discrete time goes through exactly as in the discrete time setting of [3], [21]. Below is equation (3) of [21], translated to continuous time:

$$\mathcal{W}_T(p) = \zeta\mathsf{E}_p\Big[\int_0^T \mathcal{U}(X(t))\Big]\, ds - D_T(\boldsymbol{p}\|\boldsymbol{p}^0) \tag{45}$$

Proposition 2.1 of [21] has an exact analog here: The optimizer of (45) is expressed

$$p^* = \ell_T^* p^0 \tag{46}$$

where the likelihood ratio is as before:

$$\ell^*_T(x_1^T) = \exp\Bigl(\zeta \int_0^T \mathcal{U}(x_t)\, dt - \Lambda_T(\zeta)\Bigr)$$

For example, what is the probability that $S = \int_0^T f(X(t))\, dt$ exceeds zero under p^*? Answer:

$$\mathsf{P}\{S > 0\} = \mathsf{E}_{p^0}\Bigl[\ell^*_T \mathbb{I}\{S > 0\}\Bigr]$$

where the probability is under p^*, and expectation under p^0.

The infinite-horizon average reward (12) can be expressed as the limit

$$\eta^*_\zeta = \lim_{T\to\infty} \frac{1}{T} \mathcal{W}_T(p^*_T)$$

where p^*_T is defined in (46). As in the discrete time setting, the eigenvalue Λ appearing in (28) coincides with η^*_ζ, and also the cumulative log-moment generating function:

$$\Lambda_\zeta = \lim_{T\to\infty} \frac{1}{T} \log\Bigl\{ \mathsf{E}\Bigl[\exp\Bigl(\int_0^T \zeta \mathcal{U}(X(t))\, dt\Bigr)\Bigr]\Bigr\}$$

Finally, just as in discrete time, the value $\eta^*_\zeta = \Lambda_\zeta$ is realized by a Markov process obtained through the solution of an eigenfunction problem. Let v_ζ denote a positive solution to

$$[I_{\zeta\mathcal{U}} + \mathcal{D}_0] v_\zeta = \Lambda_\zeta v_\zeta$$

One solution is given by the limit (see [14])

$$v_\zeta(x) = \lim_{T\to\infty} \mathsf{E}\Bigl[\exp\Bigl(\int_0^T [\zeta\mathcal{U}(X(t)) - \Lambda_\zeta]\, dt\Bigr) \mid X(0) = x\Bigr]$$

Exactly as in the discrete time setting, the optimal Markov process can be defined with the optimizing generator:

$$\check{\mathcal{D}}_\zeta = I^{-1}_{v_\zeta} [I_{\zeta\mathcal{U} - \Lambda_\zeta} + \mathcal{D}_0] I_{v_\zeta} \tag{47}$$

D. *Proof of Theorem 4.3*

The representation of the relative entropy rate in (36) is based on the construction of the log-likelihood ratio on a finite time horizon.

Consider two diffusions, each of which evolve according to (31) between jumps, and each with the same initial condition $x \in \mathsf{X}$. The JRF for the nominal model is denoted κ_0, and let κ denote any other jump rate. For the Markov process with JRF κ_0, functions F and G are constructed so that the following function of time is a martingale for the nominal model:

$$\ell_T = \exp\Bigl(\int_0^T G(X(s))\, ds + \sum_{i: t_i \le T} F(X(t_i^-))\Bigr) \tag{48}$$

Lemma 1.3: For any continuous JRF κ, consider the functions

$$F(x) = \log(\kappa(x)/\kappa_0(x))\,, \quad G(x) = \kappa_0(x) - \kappa(x)\,, \quad x \in \mathsf{X}\,.$$

Then the semi-group $\{P^T : T \ge 0\}$ defines the Markov process with JRF κ, where for bounded functions g,

$$P^T g\,(x) = \mathsf{E}_{\boldsymbol{p}^0}[\ell_T g(X(T)) \mid X(0) = x]\,, \quad T \ge 0\,.$$

where the expectation on the right is with respect to the Markov process with JRF κ_0.

Proof: If $g \in C^2_0$ satisfies the boundary condition (24), then the representation of the differential generator follows from the definition of the semigroup $\{P^t\}$

$$\mathcal{D} g = \frac{d}{dt} P^t g\Big|_{t=0} = \mathcal{Q}_0 g + \kappa[\bar{g} - g]$$

■

Based on Lemma 1.3, it follows that

$$D_T(\boldsymbol{p} \| \boldsymbol{p}^0) = \mathsf{E}[L_T]\,, \quad \text{with} \quad L_T = \log(\ell_T)\,,$$

and the relative entropy rate between $\mathcal{D}$ and $\mathcal{D}_0$ appearing in (11) is the limiting average

$$\begin{aligned} \mathcal{R}(\mathcal{D} \| \mathcal{D}_0) &:= \lim_{T\to\infty} \frac{1}{T} \mathsf{E}[L_T] \\ &= \lim_{T\to\infty} \frac{1}{T} \mathsf{E}\Bigl[\int_0^T G(X(s))\, ds\Bigr] \\ &\quad + \lim_{T\to\infty} \frac{1}{T} \mathsf{E}\Bigl[\sum_{i: t_i \le T} F(X(t_i^-))\Bigr] \end{aligned} \tag{49}$$

where the expectations are all with respect to $\boldsymbol{p}^0$. Applying Lemma 1.2 then establishes the formula in (37):

$$\begin{aligned} \mathcal{R}(\mathcal{D} \| \mathcal{D}_0) &= \langle \varrho, G \rangle + \langle \varrho, \kappa F \rangle \\ &= \langle \varrho, \kappa_0 - \kappa \rangle + \langle \varrho, \kappa \log(\kappa/\kappa_0) \rangle \end{aligned}$$

□

E. *Proof of Theorem 3.2*

For simplicity we suppress dependency on ζ.

To establish (29), apply (47) and simplify:

$$\begin{aligned} \check{\mathcal{D}} g &= (\zeta\mathcal{U} - \Lambda) g + v^{-1} \mathcal{D}_0 (vg) \\ &= -g v^{-1} \mathcal{D}_0 v + v^{-1} \Bigl(\rho[v' g + v g'] + \kappa_0[\bar{v}\bar{g} - vg]\Bigr) \\ &= -g v^{-1} \mathcal{D}_0 v + v^{-1} \Bigl(g \mathcal{D}_0 v + \rho v g' + \bar{v}\kappa_0[\bar{g} - g]\Bigr) \end{aligned}$$

which simplifies to $\check{\mathcal{D}} g = \rho g' + \check{\kappa}[\bar{g} - g]$, with $\check{\kappa}$ given in (29). This proves part (i).

Theorem 4.4 (ii) then implies Theorem 3.2 (ii). □