Jump Control of Probability Densities with Applications to Autonomous Vehicle Motion

Alexandre R. Mesquita, *Member, IEEE*, and João P. Hespanha, *Fellow, IEEE*

Abstract

We investigate the problem of controlling the probability density of the state of a process that is observed by the controller via a fixed but unknown scalar non-negative function of the state. The goal is to control the process so that its probability density at a point in the state space becomes proportional to the value of the function observed at that point. Our solution, inspired by bacterial chemotaxis, involves a randomized controller that switches among different deterministic modes. We show that under appropriate existence conditions, this controller guarantees convergence of the probability density to the desired function. The results can be applied to the problem of in loco optimization of a measurable signal using a team of autonomous vehicles that measure the signal but do not have access to position measurements. Alternative applications in the area of mobile robotics include deployment and environmental monitoring.

Index Terms

Piecewise-deterministic Markov processes, mobile robotics, hybrid systems

I. Introduction

This paper addresses the control of a Piecewise-Deterministic Markov Process (PDP) through the design of a stochastic supervisor that decides when switches should occur and to which mode to switch. In general, the system's state $\mathbf{x}$ cannot be measured directly and is instead observed

Manuscript received June 21, 2011. This material is based upon work supported by the Inst. for Collaborative Biotechnologies through grant DAAD19-03-D-0004 from the U.S. Army Research Office. A. R. Mesquita was partially funded by CAPES (Brazil) grant BEX 2316/05-6.

The authors are with the Center for Control, Dynamical Systems and Computation, University of California, Santa Barbara, CA 93106-9560 USA (email: mesquita@umail.ucsb.edu, hespanha@ece.ucsb.edu).

through a scalar non-negative output $\mathbf{y} = g(\mathbf{x})$, where $g(\cdot)$ is unknown to the controller. The control objective is to achieve a steady-state probability density for the state $\mathbf{x}$ that matches the unknown function $g(\cdot)$ up to a normalization factor.

We were motivated to consider this control objective by problems in the area of mobile robotics. In this type of application, $\mathbf{x}$ typically includes the position of a mobile robot that can take point measurements $\mathbf{y} = g(\mathbf{x})$ at its current location. In *deployment applications*, a group of such robots is required to distribute themselves in an environment based on the value of these measurements, e.g., the measurements may be the concentration of a chemical agent and one wants the robots to distribute themselves so that more robots will be located in areas of higher concentration of the chemical agent. In *search applications*, a group of robots is asked to find the point at which the measurement has a global maximum (or minimum), in which case one wants the probability density function of $\mathbf{x}$ to have a sharp maximum at the point x where $g(x)$ is maximum (or minimum). These applications are often referred to as "source seeking" motivated by scenarios in which the robots attempt to find the source of a chemical plume, where the concentration of the chemical exhibits a global maximum. Finally, in *monitoring applications*, one attempts to estimate the value of a spatially-defined function by keeping track of the positions of a group of robots whose spatial distribution reflects the spatially-defined function of interest (much like in deployment applications). Potential applications for this work thus include chemical plant safety, hydrothermal vent prospecting, pollution and environmental monitoring, fire or radiation monitoring, etc.

The control algorithms proposed here are motivated by the chemotactical motion of the bacterium *E. coli*. Due to its reduced size, *E. coli* is unable to perceive chemical spatial gradients by comparing measurements taken by different receptors on the cell surface. Nevertheless, this organism is still able to follow the gradient of a chemical attractant, despite the rotational diffusion that constantly changes the bacterium orientation. This is accomplished by switching between two alternate behaviors known as *run* and *tumble* [1], [2]. In the run phase, the bacterium swims with constant velocity by rotating its flagella in the counter-clockwise direction. In the tumble phase, by rotating its flagella in the clockwise direction, the bacterium spins around without changing its position and in such a way that it enters the next run phase with arbitrary orientation. Berg and Brown [1] observed that the only motion parameter that is affected by the concentration of a chemical attractant is the duration of runs. Roughly speaking, the less improvement the bacterium senses in the concentration of the attractant during the run phase, the more probable a direction change (tumble) becomes. Such a motion leads to a distribution whose peak usually coincides with the optimum of the sensed quantity, much like the search applications in mobile robotics mentioned above.

The parallel between *E. coli*'s chemotaxis and some search problems involving autonomous vehicles is remarkable: In mobile robotics, gradient information is often not directly available, either because of noisy and turbulent environments or because the vehicle size is too small to provide accurate gradient measurements, challenges also faced by *E. coli*. This bacterium also does not have access to global position information, which is analogous to the lack of position measurements that arise in applications for which inertial navigation systems are expensive, GPS is not available or not sufficiently accurate (as in underwater navigation or cave exploration), or where the vehicles are too small or weight-constrained to carry this type of equipment. These observations led us to design a biologically-inspired control algorithm for autonomous vehicles, named *optimotaxis* [3]. While mimicking chemotaxis is not a new solution to optimization problems, see e.g. [4], [5], [6], [7], [8], optimotaxis is distinct in that we are able to provide formal statements about the stationary density and the convergence to it.

In this paper, we show that the principles behind optimotaxis can be used in the much more general setting of controlling the probability density function of a PDP through the design of a stochastic supervisor that decides when switches should occur and to which mode to switch. We establish necessary and sufficient conditions under which such a controller may exist and, when these conditions hold, we provide a controller that guarantees the ergodicity of the desired invariant density. As a consequence, the probability density of the PDP converges to the desired invariant density in the Cesàro sense and results like the Law of Large Numbers apply. In addition, we provide general results that have wide application in the study of ergodicity in PDPs, beyond the specific control design problem addressed in this paper.

A substantial body of work related to the objective of controlling probability densities can be found in the literature of Markov Chain Monte Carlo (MCMC) methods [9]. These methods involve the design of a Markov chain whose stationary distribution is given by a known (but usually hard to compute) function. Samples from the Markov chain are then used to estimate integrals associated with that function. MCMC is largely used in statistical physics and in bayesian inference. According to the classification in [10], our approach can be regarded as a dynamical/hybrid MCMC type method. In particular, the *hit-and-run* method [11] resembles optimotaxis in that it also executes a piecewise linear random walk. The main difference between our approach and traditional MCMC is that the latter is a numerical method whereas the former is intended to be used in physical systems with dynamic constraints. In MCMC, for example, a trajectory may have samples discarded in order to generate a new trajectory with the desired distribution. However, this is not possible in trajectories originating from a physical system.

Our work may be related also with the field of reinforcement learning, specially with TD or Q-learning where unknown value functions are identified using only local observations of the cost function [12], and with the fields of Hidden Markov Models and particle filters, where one seeks the convergence of conditional distributions. The idea of looking at the aggregate distribution of multiple agents modeled as stochastic hybrid systems has also appeared in [13] and subsequent works.

This paper is organized as follows: the description of the problem is given in Section II; the existence and the design of controllers is discussed in Section III; Section IV presents results concerning the convergence of the probability densities of the controlled process; examples are given in Section V; conclusions and final comments are given in Section VI.

II. Problem Description

We start by briefly describing the concept of Piecewise-Deterministic Markov Processes (PDP) that is used in the paper. The reader is referred to [14] for a formal (and slightly more general) definition. In a PDP, state trajectories are right continuous with only finitely many discontinuities (*jumps*) on a finite interval. The continuous evolution of the process is described by a deterministic flow whereas the jumps occur at randomly distributed times and have random amplitudes.

We consider state variables $\mathbf{x} \in \mathbb{R}^d$ and $\mathbf{m} \in \mathbb{M}$, where $\mathbb{M}$ is a compact set. During the deterministic flows, $\mathbf{x}(t)$[1] evolves according to the vector field $x \mapsto f(x, m)$, whereas $\mathbf{m}(t)$ remains constant and only changes with jumps. For a fixed $m \in \mathbb{M}$, we denote by $\varphi_t^m x$ the

[1]We use boldface symbols to indicate random variables.

continuous flow at time t defined by the vector field $x \mapsto f(x,m)$ and starting at x at time 0. The conditional probability that at least one jump occurs between the time instants s and t, $0 < s < t$, given $\mathbf{x}(s)$ and $\mathbf{m}(s)$, is

$$1 - \exp\left(-\int_s^t \lambda(\varphi_{\tau-s}^{\mathbf{m}(s)}\mathbf{x}(s), \mathbf{m}(s))d\tau\right) \quad , \tag{1}$$

where $\lambda(x,m)$ is called the *jump rate* at $(x,m) \in \mathbb{R}^d \times \mathbb{M}$. At each jump, $\mathbf{m}$ assumes a new value governed by the *jump pdf* $T_x(\cdot,\cdot)$. Namely, if a jump occurs at time $\boldsymbol{\tau}_k$, then

$$\Pr\left\{\mathbf{m}(\boldsymbol{\tau}_k) \in B \mid \mathbf{x}^-(\boldsymbol{\tau}_k) = x, \mathbf{m}^-(\boldsymbol{\tau}_k) = m\right\} = \int_B T_x(m',m)\ \nu(dm') \quad , \tag{2}$$

where the superscript minus indicates the left limits of the respective processes, ν is a Borel probability measure on $\mathbb{M}$ and B is a Borel set. We further assume that the space $\mathbb{M}$ is a compact subset of a locally compact separable metric space and that $\operatorname{supp}\nu = \mathbb{M}$. Note that, as opposed to [14], we do not require $\mathbb{M}$ to be countable. Under this more general setting, [15] shows that the above characterization defines a strong Markov process $(\mathbf{x}(t), \mathbf{m}(t))$.

This PDP model is captured by several stochastic hybrid system models that appeared in the literature, including the stochastic hybrid models discussed in [16], or the general stochastic hybrid models introduced in [17]. Fig. 1 depicts a schematic representation of our PDP.

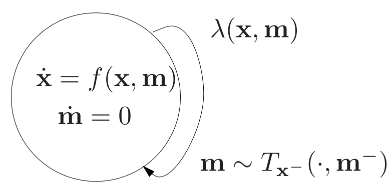

Fig. 1. Hybrid automaton for the PDP

We define $p(x,m,t)$ as the joint probability density of the state (x,m) at time t with respect to the measure $\ell \times \nu$, where ℓ denotes the Lebesgue measure in $\mathbb{R}^d$. We denote by $L^1(\mathbb{R}^d \times \mathbb{M})$ the space of real functions integrable with respect to $\ell \times \nu$. It is then true that $p \in L^1(\mathbb{R}^d \times \mathbb{M})$ and $\int_{\mathbb{R}^d \times \mathbb{M}} p(x,m,t)\ \ell(dx)\nu(dm) = 1,\ \forall t \geq 0$.

In our setting, the vector field f is given and $\mathbf{m}(t)$ should be viewed as a control variable. The controller cannot measure the state x directly; instead, an observation variable $\mathbf{y}(t) = g(\mathbf{x}(t))$

is available. In general, the function g is not known to the controller, which only has access to $\mathbf{y}(t)$.

Assuming that g is nonnegative and integrable, our objective is to design the jump rate λ and the jump pdf T_x such that a randomized controller will select $\mathbf{m}(t)$ as a function of the observations $\{\mathbf{y}(\tau); 0 \leq \tau \leq t\}$ collected up to time t so that the marginal $\int_{\mathbb{M}} p(x,m,t)\ \nu(dm)$ converges to $cg(x)$ as $t \to \infty$, where c is a normalizing constant such that cg integrates to one. We shall see later that the knowledge of the normalizing constant c is not necessary to implement the proposed control law.

In practice, g is a chosen function of some physical measurements F. For example, we can select $g(x) = Q(F(x))$, where the function $Q(\cdot)$ is a design parameter used to guarantee that $Q(F)$ is nonnegative and integrable. The function $Q(\cdot)$ may also be used to accentuate the maxima of F. For example, if the physical measurement corresponds to $F(x) = 1 - \|x\|^2$, a reasonable choice for $Q(\cdot)$ that leads to a nonnegative integrable function is

$$Q(F) = \begin{cases} F & \text{, if } F > \delta \\ \delta e^{F-\delta} & \text{, if } F \leq \delta \end{cases} , \tag{3}$$

for some $\delta > 0$. Alternatively, if one is mainly interested in the position of the maxima of $F(x)$, a possible choice for $Q(\cdot)$ is given by

$$Q(F) = F^n \ , \tag{4}$$

for some $n > 1$, provided that F^n is already nonnegative and integrable [if not one could also use Q to achieve this, as it was done in (3) above]. The well-known optimization method of simulated annealing arises from a similar objective when n is increased to infinity along consecutive iterations [18].

III. Control Design

In this section we provide a family of control laws that achieve our first objective: to make a given probability density a stationary probability density for the PDP. Only in the next section we will show convergence to this stationary density. In designing such controllers, a key auxiliary result is the generalized Fokker-Planck-Kolmogorov equation that governs the evolution of probability densities. A derivation of this equation may be found in [19, Sec. 3.4]. A more general treatment for stochastic hybrid systems is given in [20]. In the following, we use "∇_x" to denote differentiation with respect to x only. In this way, $\nabla_x \cdot fp$ denotes the divergence of fp with respect to x.

Assumption 1. *It is assumed throughout the paper that*

i. f *and* $\nabla_x f$ *are continuous functions on* $\mathbb{R}^d \times \mathbb{M}$*;*

ii. there is no finite escape time for the differential equation $\dot{x} = f(x, m)$ *and only a finite number of jumps occur in finite time intervals for the PDP* $(\mathbf{x}(t), \mathbf{m}(t))$.

□

Theorem 1. *A continuously differentiable pdf* $p(x, m, t)$ *is a pdf for* $(\mathbf{x}(t), \mathbf{m}(t))$ *if and only if it satisfies the following generalized Fokker-Planck-Kolmogorov equation:*

$$\frac{\partial p}{\partial t} + \nabla_x \cdot fp = -\lambda p + \int_{\mathbb{M}} T_x(m, m')\lambda(x, m')p(x, m', t)\nu(dm') \ . \tag{5}$$

□

Proof: Necessity follows from [20, Cor. 6]. Sufficiency follows from (4) in [20] and the fact that continuously differentiable functions separate the space of Radon measures, i.e., for two Radon measures $\mu_1 \neq \mu_2$ there exists a continuously differentiable function φ such that $\int \varphi \ d\mu_1 \neq \int \varphi \ d\mu_2$. Since PDPs do not have a diffusion component, p only needs to be continuously differentiable instead of twice continuously differentiable as in [20]. ■

When $f(x, m) = m \in \mathbb{R}^d$, equation (5) is known as the linear Boltzmann equation and has an important role in transport theory, where it models particles moving with constant velocity and colliding elastically [21]. In this case, regarding p as the density of particles, (5) has a simple intuitive interpretation: on the left-hand side we find a drift term $\nabla_x \cdot mp$ corresponding to the particles straight runs, on the right-hand side we find an absorption term $-\lambda p$ that corresponds to particles leaving the state (x, m), and an integral term corresponding to the particles jumping to the state (x, m). Equation (5) also appears in mathematical biology where it models bacterial motion [22].

Equation (5) will be used in our control design to determine a jump rate λ and a jump pdf T_x such that the joint invariant density of the process [which is obtained by setting $\partial p/\partial t = 0$ in (5)] corresponds to an invariant marginal distribution $\int_{\mathbb{M}} p(x, m, t) \ \nu(dm)$ that is proportional to $g(x)$. In fact, it will even be possible to obtain a *joint* invariant distribution $p(x,m,t)$ that is independent of m. For simplicity of presentation, in the sequel we assume that g has been scaled so that it is a probability density: $\int g(x)\ \ell(dx) = 1$. However, none of our results require this particular scaling.

A. Controller Existence and Construction

We start with an analysis that gives necessary and sufficient conditions on the vector field f for the existence of a jump control strategy that achieves the steady-state solution $p(x,m,t) = h(x,m)$, $\forall (x,m) \in \mathbb{R}^d \times \mathbb{M}$, $t > 0$, for a probability density $h(x,m)$ that integrates to 1: $\int_{\mathbb{R}^d \times \mathbb{M}} h(x,m)\ \ell(dx)\nu(dm) = 1$. We say that h is an *admissible invariant density* if there exists a jump rate λ and a jump pdf T_x such that h is an invariant density for the PDP.

Theorem 2. *Given a continuously differentiable probability density* $h(x,m) > 0$, $\forall (x,m) \in \mathbb{R}^d \times \mathbb{M}$, *with* $\nabla_x \cdot fh \in L^1(\mathbb{R}^d \times \mathbb{M})$, *a necessary and sufficient condition for* h *to be an admissible invariant density is given by*

$$\int_{\mathbb{M}} \nabla_x \cdot fh(x,m)\ \nu(dm) = 0,\ \forall x \in \mathbb{R}^d\ . \tag{6}$$

Moreover, when this condition is satisfied, the PDP has the desired invariant density h *for the uniform jump pdf* $T_x(\cdot,\cdot) \equiv 1$, *and the jump rate*

$$\lambda(x,m) = \frac{\alpha(x) - \nabla_x \cdot fh(x,m)}{h(x,m)}, \tag{7}$$

where $\alpha(x)$ *can be any function for which* λh *is nonnegative and integrable.* □

Proof: To prove necessity, assume that h is an invariant density and substitute $p(x,m,t) = h(x,m)$ in (5):

$$\nabla_x \cdot fh = -\lambda h + \int_{\mathbb{M}} T_x(m,m')\lambda(x,m')h(x,m')\nu(dm')\ . \tag{8}$$

Recall that, since $T_x(\cdot,m')$ is a pdf, $\int_{\mathbb{M}} T_x(m,m')\nu(dm) = 1$. Using this fact, condition (6) is obtained by integrating both sides of (8) on m and changing the order of integration on the right-hand side.

To prove sufficiency, we select $T_x(\cdot,\cdot) \equiv 1$ and λ as in (7), which leads to

$$\lambda h = \alpha(x) - \nabla_x \cdot fh \tag{9}$$

Provided that λh is integrable and that λ is a valid jump rate (i.e., $\lambda \geq 0$), we can replace (9) and T_x in (8) to conclude from Theorem 1 that h is indeed an invariant density for our choice of the pair (λ,T_x). One choice for the function $\alpha(x)$ that would satisfy these two conditions is $\alpha(x) = \max_{m\in\mathbb{M}} |\nabla_x \cdot fh|$. Indeed, since $\mathbb{M}$ is compact, fh is continuously differentiable and $\nabla_x \cdot fh \in L^1(\mathbb{R}^d \times \mathbb{M})$, we have that α is bounded, $\lambda \geq 0$ and $\lambda h \in L^1(\mathbb{R}^d \times \mathbb{M})$. ∎

Remark 1. It may happen that a jump rate λ satisfying (9) is not uniformly bounded, which is an issue in proving convergence to the invariant density. With $\alpha = \max_{m\in\mathbb{M}} |\nabla_x \cdot fh|$, a sufficient condition (and also necessary when (6) holds) to have $\lambda(x,m) < 2M$, $\forall(x,m)$, for some finite constant M, is $|\nabla_x \cdot fh(x,m)| \leq Mh(x,m)$, $\forall(x,m)$. □

Remark 2. The control law provided by (7) in Theorem 2 also results in the desired invariant density for a more general jump pdf: one can verify that the conclusions in the theorem hold for any jump pdf satisfying $T_x > 0$, $\int_{\mathbb{M}} T_x(m,m')\nu(dm') = 1$, $\int_{\mathbb{M}} T_x(m,m')f(x,m')\nu(dm') = 0$ and $\int_{\mathbb{M}} T_x(m,m')\nabla_x \cdot f(x,m')\nu(dm') = 0$.

Condition (6) may be restrictive on the vector field f. Since g is not known in advance, we need (6) to hold independently of g. If, however, we allow $h(x,m)$ to be arbitrary, the only vector field f that satisfies (6) for all possible densities $h(x,m)$ is $f \equiv 0$. A less restrictive condition is obtained when the desired density can be factored as $h(x,m) = \beta(m)g(x)$, $\forall(x,m)$, for some density β. In this case, the compactness of $\mathbb{M}$ and the continuity of f and of $\nabla_x f$ allow us to interchange integration and differentiation in (6) to obtain the following corollary.

Corollary 1. *Consider continuously differentiable probability densities h that can be factored as $h(x,m) = \beta(m)g(x) > 0, \forall(x,m) \in \mathbb{R}^d \times \mathbb{M}$, where $\beta > 0$ and $g > 0$ satisfy $\beta\nabla_x \cdot fg \in L^1(\mathbb{R}^d \times \mathbb{M})$. Then, a necessary and sufficient condition for all h of this form to be admissible invariant densities is given by*

$$\int_{\mathbb{M}} f(x,m)\ \beta(m)\nu(dm) = 0,\ x \in \mathbb{R}^d\ . \tag{10}$$

□

Remark 3. The existence condition (10) may be restrictive for some dynamical systems since it essentially requires the ability to “reverse” the vector field, i.e., changing the control signal from m_1 to m_2 in such a way that $f(x,m_1) = -f(x,m_2)$. This is a problem for systems with relative degree larger than zero. For example, consider the case in which $d = 2$, $\mathbb{M} = [-1, 1]$, ν is the uniform probability measure on $\mathbb{M}$ and $f(x_1, x_2, m) = [x_2 \;\, m]^T$. This PDP cannot satisfy the existence condition (6) with $h(x, m) = \beta(m)g(x)$, $\forall(x, m)$. In this case, one would be interested in achieving $\int_{\mathbb{M}} h(x, m)\nu(dm) = g(x)$ with a more general invariant density h. However, it is not clear whether that can be done using output feedback. □

B. *Output Feedback Controller*

Next we discuss whether it is possible to implement the control law (7) proposed in Theorem 2 using only information from the output. To this purpose, Corollary 1 is especially useful because the condition in (10) does not depend on the function g, which is not known in advance. We will therefore choose $h(x, m) = \beta(m)g(x)$. Without loss of generality, we set $\beta \equiv 1$, which is equivalent to redefining the reference measure to $\bar{\nu}(dm) = \beta(m)\nu(dm)$.

The uniform jump pdf $T_x(\cdot, \cdot) \equiv 1$ is trivial to implement since it does not depend on x and the controller has the freedom to select $\mathbf{m}$. Now, consider the jump rate given by (7), which we can rewrite as

$$\lambda = \eta - f \cdot \nabla_x \ln g - \nabla_x \cdot f \quad , \tag{11}$$

where $\eta(x) := \alpha(x)/g(x)$. To compute $\lambda(x, m)$, the controller needs to evaluate three terms:

- To evaluate the term $f \cdot \nabla_x \ln g$, we observe that

$$f \cdot \nabla_x \ln g(\mathbf{x}(t)) = \frac{d \ln g}{dt^+}(\mathbf{x}(t)) \quad , \tag{12}$$

  where '+' denotes the derivative from the right. Therefore, it is sufficient for the controller to have access to the time derivative of the observed output $\mathbf{y}(t) = g(\mathbf{x}(t))$ in order to evaluate this term.
- To evaluate the term $\nabla_x \cdot f$, the controller must know the vector field f and the current state x of the process. However, when $\nabla_x \cdot f$ is independent of x, state feedback is not necessary to evaluate this term.
- Regarding the term $\eta(x) = \alpha(x)/g(x)$, we have the freedom to select $\alpha(x)$ under the constraint that we keep λ nonnegative and bounded, which can be achieved if we keep $\eta \geq |f \cdot \nabla_x \ln g + \nabla_x \cdot f| = |\nabla_x \cdot fg|/g$.

In particular, when there exists some function $\phi : \mathbb{M} \to \mathbb{R}$ such that $\nabla_x \cdot f(x,m) = \phi(m), \forall (x,m)$, and a function M that satisfies $M(g) \geq \max_{m\in\mathbb{M}} |\nabla_x \cdot fg|/g$, we can use the following output feedback realization of the jump rate:

$$\lambda(\mathbf{x},\mathbf{m}) = M(\mathbf{y}) - \frac{d\ln \mathbf{y}}{dt^+} - \phi(\mathbf{m}) \ . \tag{13}$$

A slight generalization of this is used in Example A, in which an output feedback law is achieved when $\nabla_x \cdot f(x,m) = f(x,m) \cdot \nabla_x \ln \gamma(g(x)) + \gamma(g(x))\phi(m)$ for some known function γ.

Implementation of the Output Feedback Controller: Assume, for simplicity, that M is constant and that $\phi = 0$ in (13). According to (1), the probability of the process maintaining the same mode $\mathbf{m}$ in the interval $[0,t]$ is given by

$$\exp\left(-\int_0^t \lambda(\mathbf{x}(\tau),\mathbf{m}(\tau))d\tau\right) = \exp\left(-\int_0^t M - \frac{d}{d\tau}(\ln g(\mathbf{x}(\tau)))d\tau\right) = e^{-Mt}\frac{g(\mathbf{x}(t))}{g(\mathbf{x}(0))} \ . \tag{14}$$

This provides a simple and useful expression for the practical implementation of the control: Suppose that a jump happens at time $\boldsymbol{\tau}_k$. At that time pick a random variable $\mathbf{r}$ uniformly distributed in the interval $[0,1]$ and jump when the following condition holds

$$\mathbf{y}(t) \leq \mathbf{r}\ e^{M(t-\boldsymbol{\tau}_k)}\mathbf{y}(\boldsymbol{\tau}_k), \ \ t \geq \boldsymbol{\tau}_k \ . \tag{15}$$

As opposed to what (13) seems to imply, *one does not need to take derivatives of* $\ln \mathbf{y}(t)$ *to implement the jump rate*. Also, the control law is not changed if a constant scaling factor is applied to g, which is important because we cannot apply a normalizing constant to the unknown function g.

Often physical quantities propagate with spatial decay not faster than exponential, and this allows for the uniform boundedness of $\|\nabla_x \ln g\|$ and the existence of a constant M in (13). If, however, the measured quantity has a faster decay rate, it may still be possible to achieve boundedness of $\|\nabla_x \ln g\|$ by preprocessing the measurements (as explained at the end of Section II). In addition, the constant M may be identified on-line. This can be done, for example, with the following update rule:

$$\mathbf{M}(t) = \begin{cases} \epsilon\lceil \epsilon^{-1}|d(\ln \mathbf{y})/dt^+|\rceil + \epsilon, & \text{if } \mathbf{M}^-(t) < |d(\ln \mathbf{y})/dt^+| + \epsilon \\ \mathbf{M}^-(t), & \text{else} \end{cases} \tag{16}$$

for $t > 0$, $\mathbf{M}(0) = 0$ and some $\epsilon > 0$. A more elaborate adaptation could obtained by allowing $\mathbf{M}$ to depend on g. This would have the advantage of reducing the number of unnecessary jumps in some parts of the space.

IV. Ergodicity of the Controlled Process

In this section we investigate whether the above control strategy makes the probability density of the PDP converge to g as time goes to infinity. We summarize the results in this section with Theorem 3, which gives necessary and sufficient conditions for convergence.

Let $B_r(x)$ denote the open ball with radius r centered at $x \in \mathbb{R}^d$. We say that the system $\dot{x} = f(x,u)$, $u \in \mathbb{M}$, is *approximately controllable* if, for every $x_0, x_1 \in \mathbb{R}^d$ and $\epsilon_1 > 0$, there exists a time $t_1 > 0$ and a measurable control $u(t) \in \mathbb{M}$ that steers the state from $x(0) = x_0$ to $x(t_1) \in B_{\epsilon_1}(x_1)$.

Theorem 3. *Suppose that*

1. $g > 0$ is a continuously differentiable density;

2. there exists a uniformly bounded continuous function M and a constant $\epsilon > 0$ satisfying

$$|\nabla_x \cdot fg|/g + \epsilon \leq M(g);$$

3. $\nabla_x \cdot f(x,m) = \phi(m), \forall (x,m)$.

Consider the PDP $(\mathbf{x}(t), \mathbf{m}(t))$ with the output feedback control:

$$T_x(\cdot,\cdot) \equiv 1, \quad \lambda(\mathbf{x},\mathbf{m}) = M(\mathbf{y}) - \frac{d \ln \mathbf{y}}{dt^+} - \phi(\mathbf{m}) \ . \tag{17}$$

Then, $p(x,m,t) \to g(x)$ in total variation as $t \to \infty$ for all initial densities if and only if the vector field satisfies

i. $\int f(x,m)\ \nu(dm) = 0$, $\forall x \in \mathbb{R}^d$;

ii. the system $\dot{x} = f(x,u)$, $u \in \mathbb{M}$, is approximately controllable.

Moreover, the above convergence implies the following convergence of empirical averages: for every $\tau > 0$ and every ψ such that $\psi g \in L^1(\mathbb{R}^d \times \mathbb{M})$,

$$n^{-1} \sum_{k=0}^{n-1} \psi(\mathbf{x}(\tau k), \mathbf{m}(\tau k)) \to \int_{\mathbb{R}^d \times \mathbb{M}} \psi(x,m) g(x)\ \ell(dx) \nu(dm) \quad \textit{a.s.} \tag{18}$$

for all initial conditions. □

The proof of this theorem (with slightly stronger convergence) will appear later in this section. Before that, we discuss the assumptions and conclusions of the theorem.

Remark 4. We say that a set $\mathbb{F} \subset \mathbb{R}^d$ is a *positive basis* if 0 is in the (algebraic or topological) interior of the convex hull of $\mathbb{F}$. A typical case in which condition *(ii)* is satisfied is when $\{f(x,m); m \in \mathbb{M}\}$ contains a positive basis for $\mathbb{R}^d$ for all $x \in \mathbb{R}^d$ (see the Filippov-Wazewski argument in Proposition 2). If, for example, this basis is independent of x, we can always define a reference measure ν to satisfy condition *(i)*. □

In monitoring applications, the convergence of empirical averages in (18) provides the basis for a procedure to estimate g by observing the positions $\mathbf{x}_n$ of S identical vehicles performing the jump control strategy above. To achieve this, we start by partitioning the region of interest into a family of sets $\{A_i \subset \mathbb{R}^d\}$, then we sample the vehicles' positions at times $k\tau \in \{0, \tau, 2\tau, \ldots, (N-1)\tau\}$, for some $\tau > 0$, and count the frequency with which vehicles are observed in each set A_i. It turns out that this frequency provides an asymptotically correct estimate of the average value of g on the set A_i. To see why this is the case, we define

$$G_{S,N}(A_i) = \frac{1}{NS} \sum_{k=0}^{N-1} \sum_{s=0}^{S-1} 1_{A_i}(\mathbf{x}_s(k\tau)) \quad , \tag{19}$$

where 1_A denotes the indicator function of the set A. Assuming that the vehicles have mutually independent motion, we have by (18) that

$$G_{S,N}(A_i) \to G(A_i) := \int_{A_i} g(x)\ \ell(dx) \ \text{ a.s.} \tag{20}$$

as $N \to \infty$. This shows that g can be estimated by averaging the observations of the vehicles' position as in (19). The use of multiple independent agents ($S > 1$) improves the estimates according to the relation

$$\mathrm{var}(G_{S,N}) = \frac{\mathrm{var}(G_{1,N})}{S} \quad . \tag{21}$$

The following result proven in the end of the section shows that convergence is preserved if M is identified on the run using (16).

Corollary 2. *Suppose that* $\sup_{\mathbb{R}^d \times \mathbb{M}} |\nabla_x \cdot fg|/g < \infty$ *and that* $\nabla_x \cdot f = 0$*. Then, the conclusions in Theorem 3 hold when the jump rate (17) is replaced by*

$$\lambda(\mathbf{x}, \mathbf{m}) = \mathbf{M}(t) - \frac{d \ln \mathbf{y}}{dt^+} \quad , \tag{22}$$

where $\mathbf{M}(t)$ *is identified on-line using (16).*

Remark 5. In [3] it is shown that convergence of the probability density to g can also be achieved with a diffusion controller, i.e., a controller that makes use of brownian motion rather than Poisson jumps. However, the diffusion technique cannot be extended as easily to more general vector fields. Indeed, one can verify that a result similar to Theorem 3 would only be valid for vector fields that have an (complex) exponential dependence on the controlled parameters, which may be very restrictive. □

A. Elements of the Ergodic Theory for Markov Chains

Next, we present some concepts from ergodic theory that are needed to characterize the convergence of our PDP and prove Theorem 3. We consider a time-homogeneous Markov process $\mathbf{\Phi}(t)$ taking values in a locally compact separable metric space $\mathbb{Y}$ equipped with a Borel σ-algebra $\mathcal{B}$. We define the transition kernel

$$P^t(y, A) := \Pr\{\mathbf{\Phi}(t) \in A \mid \mathbf{\Phi}(0) = y\}, \quad y \in \mathbb{Y}, A \in \mathcal{B} \ . \tag{23}$$

We say that a σ-finite measure μ is an *invariant measure* for P^t if

$$\mu = \int_{\mathbb{Y}} P^t(y, \cdot)\mu(dy), \quad \forall t \geq 0 \ . \tag{24}$$

We define the *occupancy time* of the set $A \in \mathcal{B}$ as

$$\boldsymbol{\eta}_A := \int_0^\infty 1_A\{\mathbf{\Phi}(t)\} \ dt \ . \tag{25}$$

For a nontrivial σ-finite measure ψ, we say that $\mathbf{\Phi}$ is ψ*-irreducible* if, for $A \in \mathcal{B}$,

$$\psi(A) > 0 \Rightarrow \mathrm{E}[\boldsymbol{\eta}_A \mid \mathbf{\Phi}(0) = y] = \int_0^\infty P^t(y, A) \ dt > 0, \qquad \forall y \in \mathbb{Y} \ . \tag{26}$$

We say that $\mathbf{\Phi}$ is *positive* if it is ψ-irreducible and if it has an invariant measure μ satisfying $\mu(\mathbb{Y}) < \infty$. We say that $\mathbf{\Phi}$ is *aperiodic* if some sampled chain $\mathbf{\Phi}(n\tau)$ is ψ-irreducible, i.e., if there exists some $\tau > 0$ such that

$$\psi(A) > 0 \Rightarrow \sum_{n=0}^{\infty} P^{n\tau}(y, A) > 0, \ \forall y \in \mathbb{Y} \ . \tag{27}$$

A ψ-irreducible process is called *Harris recurrent* if $\psi(A) > 0$ implies that $\{\boldsymbol{\eta}_A = \infty\}$ almost surely. It is well known [23, Thm. 6.1] that aperiodic positive Harris recurrent processes are ergodic in the sense that

$$\|P^t(y, \cdot) - \mu\| \to 0 \text{ as } t \to \infty, \ \forall y \in \mathbb{Y} \ , \tag{28}$$

where μ denotes the invariant measure and the norm is the total variation norm. A set H is said *maximal absorbing* if

$$y \in H \iff \Pr\{\boldsymbol{\eta}_H = \infty \mid \boldsymbol{\Phi}(0) = y\} = 1 \ .$$

A set H is called *maximal Harris set* if H is maximal absorbing and $\boldsymbol{\Phi}$ restricted to H is Harris recurrent.

Next, we define an important continuity property used in the proof of our results. We say that $\boldsymbol{\Phi}$ is a *T-process* if, for some sampling distribution θ on $\mathbb{R}^+$,

$$R_\theta(y, A) := \int_0^\infty P^t(y, A)\ \theta(dt) \geq K(y, A),\ \ y \in \mathbb{Y}, A \in \mathcal{B} \ ,$$

where $K(\cdot, A)$ is a lower semicontinuous function for all $A \in \mathcal{B}$ and $K(y, \mathbb{Y}) > 0$ for all $y \in \mathbb{Y}$. For a ψ-irreducible T-process, we have a disjoint decomposition of the space [23, Thm. 3.4]

$$\mathbb{Y} = H \cup E \ , \tag{29}$$

where H is a maximal Harris set and E is transient in the sense that $\{\boldsymbol{\eta}_H = \infty\} \cup \{\boldsymbol{\Phi} \to \infty\}$ almost surely. We prove next that E must be open.

Lemma 1. *The set E in the decomposition (29) is an open set.*

Proof: Suppose that E is not open. Then, there is $y \in E$ such that $O \cap H \neq \emptyset$ for every neighborhood O of y . Since $y \in E$ and H is maximal absorbing, $P^t(y, E) > 0$ for all $t > 0$. Then, by Theorem [24, Thm. 9.3.2], there exists a neighborhood O of y and a distribution θ such that $R_\theta(y_0, E) > 0$ for all $y_0 \in O \cap H$. This contradicts the fact that H is maximal absorbing. Therefore, E must be open. ■

B. Ergodicity for the PDP

In this section we derive some new results regarding the ergodicity of invariant measures of PDPs and prove Theorem 3. While some ergodicity results specific for PDPs may be found in the literature (see e.g. [25]), those rely mostly on Foster-Lyapunov criteria and do not appear to be suited for the purposes of this paper, since we try to prove ergodicity for general vector fields f. On the other hand, our task of proving ergodicity is made somewhat easier since we know, by design, that an invariant measure exists.

Let us call *jump Markov chain* a new PDP obtained from the original one by replacing the vector field f by $f(x,m)=0$, $\forall(x,m)$. We say that the jump Markov chain is *mode-irreducible* if, for each initial $(x,m)\in\mathbb{R}^d\times\mathbb{M}$ and any set B with $\nu(B)>$0, there is a positive probability that $\{x\}\times B$ will eventually be reached from (x,m). The following assumption will be needed in the results of this section.

Assumption 2. *i. the jump Markov chain is mode-irreducible.*

ii. $\lambda(x,m)$ *is a bounded continuous function on* $\mathbb{R}^d\times\mathbb{M}$ *and, for any bounded and continuous* ψ*, the map*

$$(x,m)\mapsto\int_{\mathbb{M}}T_x(m',m)\psi(x,m')\ \nu(dm') \tag{30}$$

is continuous.

iii. $\int f(x,m)\ \nu(dm)=0$*;*

iv. the system $\dot{x}=f(x,u)$*,* $u\in\mathbb{M}$*, is approximately controllable.*

□

We denote by P^t the transition kernel of the PDP $(\mathbf{x}(t),\mathbf{m}(t))$.

Proposition 1. *Suppose that Assumption 2 (i)-(ii) holds. Let* $\bar{m}:\mathbb{R}^+\to\mathbb{M}$ *be a piecewise constant function with finitely many jumps and let* $\bar{x}(t)$ *be the solution to* $\dot{\bar{x}}(t)=f(\bar{x}(t),\bar{m}(t))$ *with initial condition* $\bar{x}_0$*. Then, given the initial condition* $(\bar{x}_0,\bar{m}(0))$ *and any* $\epsilon_1,t_1>0$*, the PDP* $(\mathbf{x}(t),\mathbf{m}(t))$ *visits the ball of radius* ϵ_1 *centered at* $(\bar{x}(t_1),\bar{m}(t_1))$ *with positive probability at time* $t=t_1$*, i.e.,*

$$P^{t_1}\big((\bar{x}(0),\bar{m}(0)),B_{\epsilon_1}(\bar{x}(t_1),\bar{m}(t_1))\big)>0,\ \ \forall t_1,\epsilon_1>0. \tag{31}$$

□

Proof: By Assumption 2 *(i)*, given a time $t_1>0$ and $\epsilon_0>0$, there exists $m(t)$ satisfying $m(t)\in\operatorname{supp}T_{\bar{x}(t)}(\cdot,m^-(t))$ and $m(t)=\bar{m}(t)$ on $[0,t_1]\backslash S$, where S has Lebesgue measure ϵ_0. Thus, if $\dot{x}(t)=f(x(t),m(t))$ and $x(0)=\bar{x}_0$, the assumption of continuity of $\nabla_x f$ and of no finite scape time implies have that $\|x(t_1)-\bar{x}(t_1)\|<\kappa\epsilon_0$ for some constant $\kappa>0$. On the other hand, the smoothness of f and the irreducibility and continuity assumptions in Assumption 2 *(i)-(ii)* imply that $(\mathbf{x}(t_1),\mathbf{m}(t_1))$ is found in any neighborhood of $(x(t_1),m(t_1))$ with positive probability. Combining the two facts, we have the result in the proposition. ■

Let $\overline{\text{co}}(A)$ denote the closure of the convex hull of the set A.

Proposition 2. *Suppose that Assumption (i)-(ii) holds and let* $\hat{x}(t)$ *be a solution to the differential inclusion* $\dot{\hat{x}} \in \overline{\text{co}}\{f(\hat{x}, m), m \in \mathbb{M}\}$ *with initial condition* x_0*. Then, given* $\epsilon_1, t_1 > 0$ *and* $m_0, m_1 \in \mathbb{M}$*, the PDP* $(\mathbf{x}(t), \mathbf{m}(t))$ *with initial condition* (x_0, m_0) *visits the ball of radius* ϵ_1 *centered at* $(\hat{x}(t_1), m_1)$ *with positive probability at time* $t = t_1$*, i.e.,*

$$P^t\big((x_0, m_0), B_{\epsilon_1}(\hat{x}(t), m_1)\big) > 0,\ \forall t, \epsilon_1 > 0 \tag{32}$$

and for all $m_0, m_1 \in \mathbb{M}$*. As a consequence, under Assumption (i)-(ii), approximate controllability is equivalent to* $\ell \times \nu$*-irreducibility.* □

Proof: Let $x_u(t)$ denote the solution to $\dot{x}_u(t) = f(x_u(t), u(t))$ for the initial condition x_0 and some control $u(t)$. By the continuity of $\nabla_x f$ and the assumption of no finite scape time (Assumption 1), we can apply the Filippov-Wazewski theorem [26, Thm 10.4.3] to conclude that, given $t_1, \epsilon_0 > 0$, there exists a measurable control $u(t) \in \mathbb{M}$ such that $\|x_u(t_1) - \hat{x}(t_1)\| < \epsilon_0$. Under Assumption 1 *(i)*, we can apply Theorems 2.20 and 2.24 of [27] to conclude that there exists a piecewise-constant control $m(t) \in \mathbb{M}$ with finitely many jumps that approximates the measurable control $u(t)$ in the sense that $\|x_u(t_1) - x_m(t_1)\| < \epsilon_0$. Thus, by Proposition 1, we conclude that $P^{t_1}((x_0, m(0)), B_{\epsilon_1}(\hat{x}(t_1), m(t_1))) > 0$ for any $\epsilon_1 > 0$. As in the proof of Proposition 1, this holds for arbitrary initial and final modes $m(0)$ and $m(t_1)$ since the PDP is jump-irreducible and m may take arbitrarily small time on those states. ■

Similarly to [24, Chap. 7], we establish a link between controllability and irreducible T-processes in the next proposition.

Proposition 3. *Under Assumption 2 (i)-(iv), the PDP* $(\mathbf{x}(t), \mathbf{m}(t))$ *is an aperiodic* $\ell \times \nu$*-irreducible T-process.* □

Proof: By Proposition 2, $\ell \times \nu$-irreducibility is equivalent to the controllability condition *(iii)*. From condition *(iv)*, we have that $0 \in \overline{\text{co}}\{f(x, m); m \in \mathbb{M}\}$ and Proposition 2 implies that $P^t((x_0, m_0), B_{\epsilon_0}(x_0, m_0)) > 0$ for all $t > 0$ and $\epsilon_0 > 0$. This implies aperiodicity of the $\ell \times \nu$-irreducible process since trajectories starting on any open set return to the set at any time with positive probability. By [14, Thm. 27.6], Assumption 2 (*ii*) implies that the PDP has the (weak) Feller property, i.e., the function $\xi \mapsto \int_{\mathbb{R}^d \times \mathbb{M}} \psi(y) P^t(\xi, dy)$ is continuous for all bounded continuous functions ψ and all $t > 0$. Given the Feller property, the $\ell \times \nu$-irreducibility and the fact that $\text{supp}(\ell \times \nu) = \mathbb{R}^d \times \mathbb{M}$ has non-empty interior, we can use [24, Thm. 6.0.1] to conclude that the PDP is a T-process. ■

Proposition 4. *Suppose that Assumption 2 (i)-(iv) holds and that the PDP* $(\mathbf{x}(t), \mathbf{m}(t))$ *admits an invariant probability density* $h(x, m) > 0$*. Then, the PDP is an aperiodic positive Harris recurrent process and convergence to the invariant measure in total variation holds as in (28).* □

Proof: Let μ denote the invariant measure corresponding to h. By Proposition 3, the PDP is an aperiodic $\ell \times \nu$-irreducible T-process. Therefore, the space $\mathbb{R}^d \times \mathbb{M}$ admits a decomposition into a maximal Harris set H with invariant measure μ and a transient set E as in [23, Thm. 3.4]. Since $\mu(H) = 1$ and $h(x, m) > 0$, we must have $\ell \times \nu(E) = 0$. However, E is an open set by Lemma 1. This implies that $E = \emptyset$ and therefore the PDP is an aperiodic positive Harris recurrent process. By [23, Thm. 6.1], we have convergence as in (28). ■

Proof of Theorem 3: (Necessity) The necessity of condition *(i)* follows from Corollary 1. To see the necessity of condition *(ii)*, note that the convergence of $p(x, m, t)$ to $g(x)$ implies that the process is μ-irreducible, where $\mu(dx, dm) = g(x)\ell(dx)\nu(dm)$, which implies the controllability condition since $g > 0$.

(Sufficiency) It follows from Theorem 2 and Corollary 1 that g is an invariant density for the pair (λ, T_x) presented. To prove convergence, we show that Assumption 2 holds and apply Proposition 4. The inequality $|\nabla_x \cdot fg|/g + \epsilon < M$ implies that $\lambda = M - (\nabla_x \cdot fg)/g \geq \epsilon$ and that λ is uniformly bounded. Therefore, Assumption 2 *(i)* holds since $\lambda \geq \epsilon$ and a uniform jump distribution imply mode-irreducibility. Since fg is continuously differentiable in x, we have that λ is continuous and, therefore, Assumption 2 *(ii)* holds. Assumption 2 *(iii)-(iv)* follows from conditions *(i)* and *(ii)*. Therefore, we have that the process is aperiodic positive Harris recurrent and convergence in total variation holds. Clearly, the same convergence result as in (28) must hold for the kernel $P^{k\tau}$. This implies that $(\mathbf{x}(k\tau), \mathbf{m}(k\tau))$ is positive Harris for all $\tau > 0$. Then, the convergence of the empirical averages for all initial conditions follows from [24, Thm. 17.0.1]. ■

Proof of Corollary 2: We consider the Markov process formed by $(\mathbf{x}(t), \mathbf{m}(t), \mathbf{M}(t))$. Let

$\bar{M} = \epsilon\lceil \epsilon^{-1} \sup_{\mathbb{R}^d\times\mathbb{M}} |\nabla_x \cdot fg|/g\rceil + \epsilon$. From (16), $\mathbf{M}(t)$ increases by at least ϵ at every update. Thus, $\mathbf{M}(t)$ must achieve a limit $\mathbf{M}_0 \le \bar{M}$ in finite time almost surely. Suppose $\mathbf{M}_0 \le \bar{M} - \epsilon$ and let $C_0 = \{(x, m) \in \mathbb{R}^d \times \mathbb{M} : |f \cdot \nabla_x \ln g| + \epsilon \le \mathbf{M}_0\}$. This definition implies that $(\mathbf{x}(t), \mathbf{m}(t)) \in C_0$ for all time. Let g_0 be a probability density such that $g_0 = g$ on C_0 and $|f \cdot \nabla_x \ln g_0| + \epsilon \le \mathbf{M}_0$. Since $\lambda \ge 0$ on C_0, we can apply Theorem 3 to conclude that $p(x, v, t)$ converges to g_0. But, since $\mathbf{M}_0 \le \bar{M} - \epsilon$,

$$\int_{C_0^c} g(x)\ell(dx)\nu(dm) \ge \int_{\{|f\cdot\nabla_x \ln g|+\epsilon>\bar{M}-\epsilon\}} g(x)\ell(dx)\nu(dm) > 0 \ , \tag{33}$$

where the last inequality follows from the continuity of $f \cdot \nabla_x \ln g$. This contradicts the convergence of $p(x, m, t)$ to $g_0(x)$ since $\int_{C_0} g_0(x)\ell(dx)\nu(dm) < 1$. Therefore, $\mathbf{M}(t)$ achieves the limit $\mathbf{M}_0 = \bar{M}$ in finite time almost surely. From the proof of Theorem 3, this is the same to say that $(\mathbf{x}(t), \mathbf{m}(t), \mathbf{M}(t))$ reaches the Harris recurrent set $\mathbb{R}^d \times \mathbb{M} \times \{\bar{M}\}$ in finite time almost surely from any initial condition. Using the strong Markov property as in [24, Prop. 9.1.1], this implies that the process is Harris recurrent and the proof proceeds as in the proof of Theorem 3. ■

V. Examples

In this section we present applications of our main result to three systems characterized by different dynamics. The first dynamics are heavily inspired by the tumble and run motion of *E. coli* and correspond to a vehicle that either moves in a straight line or rotates in place. The second is a Reeds-Shepp car [28], which has turning constraints, but can reverse its direction of motion instantaneously. The third dynamics corresponds to a vehicle that is controlled through attraction/repulsion by one of three beacons in the plane.

A. Optimotaxis

Optimotaxis was introduced in [3] as a solution to an in loco optimization problem with point measurements only. We consider vehicles moving with position $\mathbf{x} \in \mathbb{R}^d$ and velocity $\mathbf{v}\rho(\mathbf{x})$, where $\mathbf{v}$ belongs to the unit sphere $\mathbb{M} = \mathbb{S}^d$ and the uniformly bounded function $\rho(x)$ is the space-dependent velocity amplitude. The reference measure ν is the normalized surface measure on the sphere. In this case, the mode is represented by $\mathbf{v}$ and we have $f(x, v) = v\rho(x)$. Our objective is to make the probability density of the vehicles position converge to an observed function g and then have an external observer that can measure the vehicles position to collect information about g.

In [3], we were forced to consider a constant velocity amplitude ρ, but now we can allow the velocity amplitude ρ to depend on x through the output $g(x)$. This modification is important since it is generally advantageous to move fast through regions where g is small and slowly through regions where g is large. This idea is pursued further in [29].

Here we apply Theorem 3 with a small modification. The assumption in the theorem that $\nabla_x \cdot f$ be independent of x is only needed to make sure that λ in (11) can be implemented with output feedback. Although this assumption is not satisfied in our example, we can still manipulate (11) to obtain the following output feedback implementation:

$$\lambda(x,v) = \eta\rho - v\rho \cdot \nabla_x \ln \rho g \ \ , \tag{34}$$

where $\eta > \|\nabla_x \ln \rho g\|$. Because the divergence of f is nonzero, we must alter the implementation rule (15) to

$$\rho(\mathbf{x}(t))\mathbf{y}(t) \leq \mathbf{r}e^{\eta(t-\tau_k)}\rho(\mathbf{x}(\tau_k))\mathbf{y}(\tau_k), \ \ t \geq \tau_k \ \ . \tag{35}$$

Since $\{f(x,v); v \in \mathbb{M}\}$ is a positive basis for all x provided that $\rho(x) > 0, \forall x$ (see Remark 4), we conclude convergence of our controlled process as in Theorem 3.

Next, we present numerical experiments to illustrate the proposed optimization procedure. The desired stationary density is taken to be $g(x) = cF^n(x)$, where F are the physical measurements, c is a normalizing constant and n is an integer.

The ability of optimotaxis to localize the global maximum is stressed in Fig. 2. We observe a swarm of agents that starts from the upper left corner (I), initially clusters around a local maximum (II) and then progressively migrates to the global maximum (III,IV). We notice that the center of mass of the swarm goes straight through the local maximum to the global one. When the equilibrium is reached, most agents concentrate in a neighborhood of the global maximum, while a few remain near the local maximum as one should expect. The proportion of agents at each maximum reflects the values of g as expressed by (20).

To quantify the convergence of the positions of the agents to the desired distribution g, we compute the correlation coefficient between the vectors $[G(A_i)]_i$ and $[G_{S,N}(A_i)]_i$ in (20).

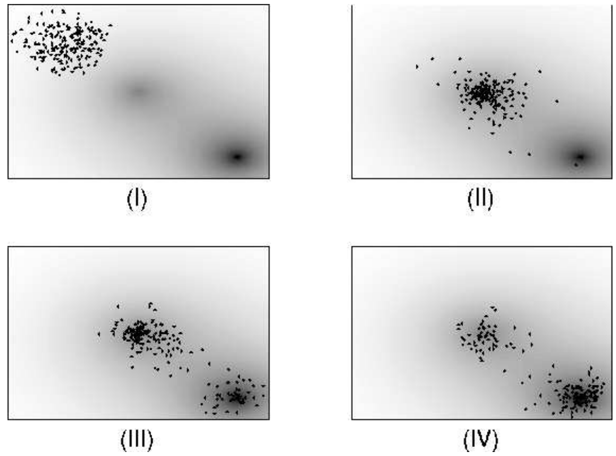

Fig. 2. Different stages of optimotaxis in the presence of two maxima. Black dots represent agents position whereas the background intensity represents the signal intensity. $F(x) = 0.4e^{-\|x\|} + 0.6e^{-\|x-[1.5\ -1.5]'\|}$, $g(x) = F^n(x)$ with $n = 10$, $\rho(x) \equiv 1$ and η chosen as in (16).

This coefficient was calculated using a space grid with resolution 0.068 and its time evolution appears Fig. 3. Also included in Fig. 3 is the evolution of the correlation coefficient when the measurements are quantized and when exogenous noise is added. For the quantized case, we used a quantized version of the desired density g to calculate the coefficient. Interestingly, the addition of noise does not seem to affect considerably the speed of convergence. Nevertheless, the residual error is greater due to the fact that the observed stationary density is not exactly equal to g. On the other hand, quantization has a negative impact on convergence time.

Many factors may affect the convergence speed. In optimization applications with $g = F^n$, we have studied in [3] the influence of the parameter n on the convergence. The results suggest that there exists an optimal choice of n that maximizes the speed of convergence. The velocity amplitude $\rho(x)$ is another design parameter that influences convergence speed. In Fig. 4, we observe that the transient response for $\rho \equiv 25$ is initially faster when compared to the transient response for $\rho \equiv 10$, but it becomes ultimately slower after a level of 70% of correlation is reached. In the same figure we see that one can speed up convergence by adjusting $\rho(x)$ to be large when $g(x)$ is small and to be small when $g(x)$ is large. The use of a space dependent velocity amplitude $\rho(x)$ is further studied in [29].

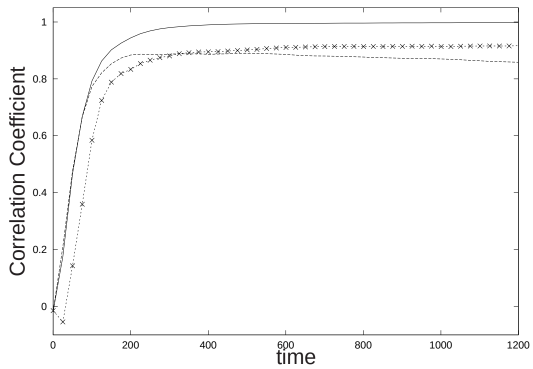

Fig. 3. Evolution of the coefficient of correlation for: the noiseless case (solid), the quantized measurements case (cross), and the exogenous noise case (dashed). The number of quantization levels is 64. The noise added to $\dot{v}$ is white Gaussian with standard deviation 10^{-2} along each axis. $N = 100$ agents were uniformly deployed in the rectangle $[-2.5, -1.5] \times [1.5, 2.5] \times \mathbb{M}$ and simulated with sampling time 1. Refer to Fig. 2 for more details.

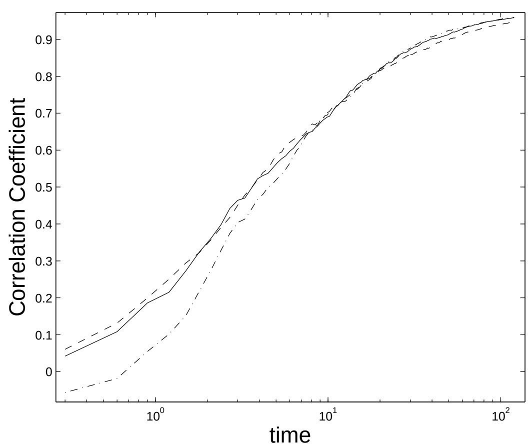

Fig. 4. Correlation coefficient for the cases $\rho \equiv 25$ (dashed), $\rho \equiv 10$ (dash-dotted) and $\rho = 25 \tanh(g^{-2}/25)$ (solid) with $n = 1$ and other simulation details as in Figs. 2 and 3.

Chemotaxis and Optimotaxis: Chemotaxis in the bacterium *E. coli* can be seen as an example of how the jump control of probability densities is used for the optimal distribution of individuals. The run and tumble behavior discussed in the introduction can be cast into our optimotaxis framework for some specific expressions for λ and T_x. It is remarkable that the expression for λ in (34) obtained in the optimotaxis example is an affine function of $d(\ln \mathbf{y})/dt$, which coincides with simple biochemical models for the tumbling [jump] rate of the *E. coli*;

see, for instance, [30, Sections 6.1 and 8.3] or Alt [2, Equation 4.8]. The latter author proposed the existence of a chemical activator for the locomotion mechanism such that a tumble would take place each time the concentration of this activator becomes smaller than a certain value. The concentration of this activator would jump to a high value at tumbles and decrease at a rate corresponding to η in (34). A receptor-sensor mechanism would then regulate the additional generation of the activator [this corresponds to the term $v \cdot \nabla \ln g(x)$ in (34)], which would modulate the run length.

Though the use of tumble and run in optimotaxis was inspired by chemotaxis, one would not necessarily expect that our choice of the jump rate would lead to control laws that resemble the biochemical models in bacteria. More precisely, we have borrowed the PDP structure from chemotaxis, but the parameters λ and T_x were designed independently of it. It turns out that, if we make the natural assumption that the jump pdf is uniformly distributed, the jump rate given by (34) is the unique jump rate that achieves a stationary density g (this can be seen from the Fokker-Planck-Kolmogorov equation). This suggests that also bacterial chemotaxis is aimed at achieving a stationary density that is a function of the measured profile of chemical attractant. As a consequence of this fact, our control law can be used to analyze the bacterial motion and to predict what stationary distribution is aimed by the bacteria.

Let us imagine that bacteria are performing optimotaxis as it is described in this paper, let $p(x, v, t)$ be the spatial density of bacteria and let $g(x)$ be some function related to the concentration of nutrients at point x. Suppose also that the bacteria are in a static environment like a chemostat, which would maintain the level of nutrients constant in time, or that the consumption of nutrients happens in a timescale that is much slower than chemotaxis. Under these conditions, we will show that we can translate the objective of $p(x, v, t)$ converging in total variation to g as the minimization of a biologically relevant quantity. From [31], convergence of p to g implies that

$$H(t) = -\int_{\mathsf{X}} \int_{\mathbb{M}} p(x, v, t) \ln \left(\frac{1}{2} + \frac{1}{2} \frac{g(x)}{p(x, v, t)} \right) \ \ell(dx)\nu(dv) \to 0 \ , \tag{36}$$

where $H(t)$ is the Kullback-Leibler divergence between $p(x, v, t)$ and the convex combination $1/2\ g(x) + 1/2\ p(x, v, t)$. Since $H(t) \geq 0, \forall t$, and is equal to zero if and only if $g(x) = p(x, v, t)$ a.e., one can regard $H(t)$ as a cost functional that is being minimized by bacterial chemotaxis (and, in fact, also by optimotaxis). More specifically, we notice that what is being maximized is the expected value of an increasing concave function of g/p, which is a ratio that measures the concentration of nutrients per density of organisms. Thus, what is being maximized here is not the probability of a bacterium being at the point of maximum concentration of nutrients, but the average amount of nutrients a bacterium has access to when interacting with many others of its kind, which is a biologically meaningful cost for the population of bacteria as a whole. Interestingly, this effect is achieved as a result of an individualistic behavior (without direct interaction among the bacteria), which suggests that it could arise as an evolutionary equilibrium.

Since our conclusions are based on a simplistic model for chemotaxis, further investigation is necessary. Nevertheless, our analysis suggests an optimal cooperative aspect of chemotaxis that is original to the best of our knowledge. The idea of cooperation among bacteria is corroborated by the phenomenon of chemotactic signaling, according to which bacteria may cooperate by emitting attractants or repellents to indicate to others the presence or scarcity of nutrients respectively [32]. Similar conclusions regarding an adaptation to the spatial density of preys may be drawn for predators following the work in [33]. The optimality of chemotaxis in the sense of tracking chemical gradients was investigated in [34].

B. Example 2

We now consider optimotaxis when vehicles are subject to turning constraints but are still able to immediately change between forward and backward motion. More precisely, the dynamics of the vehicle is given by

$$f(x,v) = \begin{bmatrix} v_1 \cos\theta \\ v_1 \sin\theta \\ \omega \end{bmatrix} , \tag{37}$$

where $x = [x_1 \ x_2 \ \theta]'$, $v = [v_1 \ \omega]' \in \mathbb{M} = \{-v_0, 0, v_0\} \times \{-\omega_0, 0, \omega_0\}$ and ν is the uniform probability density over $\mathbb{M}$. This kind of vehicle is referred to in the literature as the Reeds-Shepp car [28].

The vector field in (37) satisfies condition (*i*) of Theorem 3 and, even though $\{f(x,v);\ v \in \mathbb{M}\}$ does not contain a positive basis for $\mathbb{R}^3$ as in the previous example, it is still an easy exercise to verify condition (*ii*) in the Theorem 3 by constructing trajectories between any two points in $\mathbb{R}^3$.

This controllability condition would hold true even if zero linear velocity was not allowed. Indeed, we know that it is always possible to steer a Dubins' vehicle between any two configurations [35]. Since a Dubins' vehicle is a special case of the Reeds-Shepp car in which only positive linear velocity is allowed, we have that it is also possible to steer a Reeds-Shepp car between any two configurations (states). Hence, we can use λ and T_x from Theorem 3 to make the process pdf converge to the invariant density g. More precisely, $\lambda = \eta - f \cdot \nabla_x \ln g$ and $T_x \equiv 1$.

Figure 5 illustrates how the empirical distribution indeed converges to the desired density. To evaluate the effect of turning constraints, we also plot the unconstrained case discussed in the previous example. The figure shows that convergence is only slightly slower compared to the unconstrained case when $\omega_0 = 0.3$, but there is a strong dependence in the turning speed as shown when this speed is decreased by a factor of 2. It is worth to mention that in the case for which 0 linear velocity is not allowed convergence is only slightly slower than in the present case.

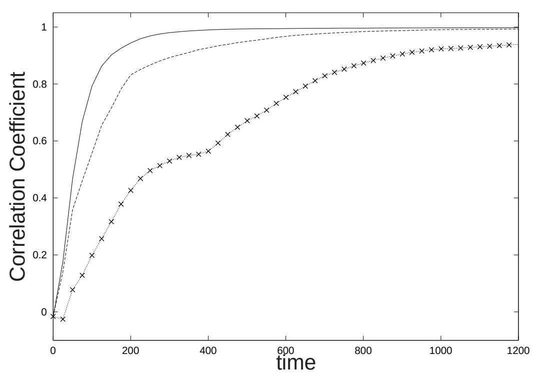

Fig. 5. Evolution of the coefficient of correlation for the unconstrained turning case (solid), for the constrained turning case with $v_0 = 1$ and $\omega_0 = 0.3$ (dots) and for the constrained turning case with $v_0 = 1$ and $\omega_0 = 0.15$ (cross). The simulation setting is the same as of Figs. 2 and 3.

C. Example 3

In this example, vehicles make use of three beacons in order to navigate. In particular, vehicles always move along straight lines towards or away from one of the beacons located at positions $a, b, c \in \mathbb{R}^2$. Let $d = 2$, $\mathbb{M} = \{a, b, c\} \times \{-1, 1\}$, where a, b, c are not in the same line, and ν is the uniform probability distribution over $\mathbb{M}$. We take $f(x, m)$ to be $f = m_2(x - m_1)$,

$x \in \mathbb{R}^2$, $m_1 \in \{a, b, c\}$, $m_2 \in \{-1, 1\}$. Thus, a, b and c are three points in the plane that may be either stable or unstable nodes (depending on the sign of m_2). This is an example for which the divergence is not zero. According to Theorem 3, we choose

$$\lambda = M - f \cdot \nabla \ln g - 2m_2 \ , \tag{38}$$

for some M sufficiently large. Note that f satisfies the hypotheses of Theorem 3: since a, b and c are not aligned, $\{f(x, m); m \in \mathbb{M}\}$ is a positive basis (see Remark 4). The class of reachable densities includes those for which $x\|\nabla \ln g\|$ is uniformly bounded, which includes all densities with polynomial decay. We note that a uniform T_x is not the only one that achieves the desired density for such a λ. For example, as noted in Remark 2, it is possible to choose T_x such that

$$T_x(m, m') = \frac{1}{4} 1_{\mathbb{M} - \{m'_1 \times \{-1,1\}\}}(m) \ . \tag{39}$$

This jump pdf is such that jumps to the flows with the same fixed point are not allowed. Yet, since λT_x still defines a mode-irreducible Markov chain, we can apply Proposition 4 to conclude convergence to the invariant density g.

VI. Conclusion

A solution to the problem of controlling the probability densities of a process was provided. Our solution, which involves a randomized controller that switches among different deterministic modes, is applicable when the observation process is a fixed but unknown function of the state. Necessary and sufficient conditions were derived to determine when such a controller can enforce a given stationary density for the process. We also provide conditions under which the probability density of the process converges to the observation function. We discussed potential applications of this theory to the area of mobile robotics, where it can be used to solve problems including search, deployment and monitoring.

One challenge to be addressed in the future is the development design tools for systems with relative degree higher than or equal to one, as discussed in Remark 3. A second important problem is to define convergence rates in a manner that is useful for both analysis and design. Possible frameworks include the use of Lyapunov techniques or the theory of Large Deviations as in [29]. In addition, the authors believe it would be beneficial to explore new applications for their method in the existing large domain of applications for Markov Chain Monte Carlo methods.

ACKNOWLEGEMENTS

The authors would like to thank K. J. Åström for his contribution to our early work [3], which motivated this paper.

References

[1] H. Berg and D. Brown, "Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking," *Nature*, vol. 239, no. 5374, pp. 500–504, Oct. 1972.

[2] W. Alt, "Biased random walk models for chemotaxis and related diffusion approximations." *J Math Biol*, vol. 9, no. 2, pp. 147–177, Apr. 1980.

[3] A. Mesquita, J. Hespanha, and K. Åström, "Optimotaxis: A stochastic multi-agent optimization procedure with point measurements," in *Hybrid Systems: Computation and Control*, M. Egerstedt and B. Mishra, Eds. Berlin: Springer-Verlag, Mar. 2008, no. 4981, pp. 358–371, available at http://www.ece.ucsb.edu/~hespanha/published.

[4] D. Hoskins, "Least action approach to collective behavior," in *Proc. SPIE Microrobotics and Micromechanical Systems*, L. E. Parker, Ed., vol. 2593, Dec. 1995, pp. 108–120.

[5] M. Vergassola, E. Villermaux, and B. Shraiman, "Infotaxis as a strategy for searching without gradients," *Nature*, vol. 445, no. 7126, pp. 406–409, 2007.

[6] T. Ferrée and S. Lockery, "Computational rules for chemotaxis in the nematode *C. Elegans*," *Journal of Computational Neuroscience*, vol. 6, pp. 263–277, 1999.

[7] A. Dhariwal, G. Sukhatme, and A. Requicha, "Bacterium-inspired robots for environmental monitoring," in *IEEE International Conference on Robotics and Automation*. New Orleans, Louisiana: IEEE, Apr 2004, pp. 1436–1443. [Online]. Available: http://cres.usc.edu/cgi-bin/print_pub_details.pl?pubid=375

[8] A. Linhares, "Synthesizing a predatory search strategy for VLSI layouts," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 2, pp. 147–152, 1999.

[9] W. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC, 1996.

[10] R. Neal, *Bayesian Learning for Neural Networks*. Springer, 1996.

[11] S. Vempala, "Geometric random walks: A survey," *MSRI volume on Combinatorial and Computational Geometry*, 2005.

[12] D. Bertsekas and J. Tsitsiklis, *Neuro-Dynamic Programming (Optimization and Neural Computation Series, 3)*. Athena Scientific, May, 1996.

[13] R. Malhamé and C. Chong, "Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system," *IEEE Transations on Automatic Control*, vol. 30, no. 9, pp. 854–860, September 1985.

[14] M. Davis, *Markov models and optimization*, ser. Monographs on statistics and applied probability. London, UK: Chapman & Hall, 1993.

[15] M. Jacobsen, *Point Process Theory and Applications: Marked Point and Piecewise Deterministic Processes*. Birkhauser, 2006.

[16] J. Hespanha, "Modeling and analysis of stochastic hybrid systems," *IEE Proc — Control Theory & Applications*, Special Issue on Hybrid Systems, vol. 153, no. 5, pp. 520–535, 2007.

[17] M. Bujorianu and J. Lygeros, "General stochastic hybrid systems: Modelling and optimal control," in *43rd IEEE Conference on Decision and Control, 2004. CDC*, vol. 2, 2004.

[18] P. van Laarhoven and E. Aarts, "Simulated annealing: theory and applications," *Mathematics and Its Applications, D. Reidel, Dordrecht*, 1987.

[19] C. Gardiner, *Handbook of stochastic methods*. Springer Berlin, 1985.

[20] J. Bect, "A unifying formulation of the Fokker-Planck-Kolmogorov equation for general stochastic hybrid systems," *Nonlinear Analysis: Hybrid Systems*, 2009.

[21] H. Kaper, C. Lekkerkerker, and J. Hejtmanek, *Spectral Methods in Linear Transport Theory*. Birkhauser Verlag, 1982.

[22] H. Othmer, S. Dunbar, and W. Alt, "Models of dispersal in biological systems," *Journal of Mathematical Biology*, vol. 26, no. 3, pp. 263–298, 1988.

[23] S. Meyn and R. Tweedie, "Stability of Markovian processes II: Continuous-time processes and sampled chains," *Advances in Applied Probability*, pp. 487–517, 1993.

[24] ——, *Markov chains and stochastic stability*. Cambridge Univ Pr, 2009.

[25] O. L. V. Costa and F. Dufour, "Stability and ergodicity of piecewise deterministic markov processes," *SIAM J. Control Optim.*, vol. 47, no. 2, pp. 1053–1077, 2008.

[26] J. Aubin and H. Frankowska, "Differential inclusions," *Set-Valued Analysis*, pp. 1–27, 2009.

[27] K. Grasse and H. Sussmann, "Global controllability by nice controls," *Nonlinear controllability and optimal control*, pp. 33–79, 1990.

[28] H. Sussmann and G. Tang, "Shortest paths for the Reeds-Shepp car: A worked out example of the use of geometric techniques in nonlinear optimal control," *SYCON report*, vol. 9110, 1991.

[29] A. Mesquita and J. Hespanha, "Construction of Lyapunov Functions for Piecewise Deterministic Processes," in *49th IEEE Conference on Decision and Control*, 2010, available at http://www.ece.ucsb.edu/~hespanha/published.

[30] R. Erban and H. Othmer, "From individual to collective behavior in bacterial chemotaxis," *SIAM Journal on Applied Mathematics*, vol. 65, no. 2, pp. 361–391, 2004.

[31] J. Lin, "Divergence measures based on shannon entropy," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 145–151, Jan. 1991.

[32] E. Ben-Jacob, I. Cohen, and H. Levine, "Cooperative self-organization of microorganisms," *Advances in Physics*, vol. 49, no. 4, pp. 395–554, 2000.

[33] P. Kareiva and G. Odell, "Swarms of Predators Exhibit "Preytaxis" if Individual Predators Use Area-Restricted Search," *American Naturalist*, vol. 130, no. 2, p. 233, 1987.

[34] S. Strong, B. Freedman, W. Bialek, and R. Koberle, "Adaptation and optimal chemotactic strategy for E. coli," *Physical Review E*, vol. 57, no. 4, pp. 4604–4617, 1998.

[35] L. Dubins, "On curves of minimal length with a constraint on average curvature, and with prescribed initial and terminal positions and tangents," *American Journal of Mathematics*, vol. 79, no. 3, pp. 497–516, 1957.