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 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 controllability 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 use point measurements of 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

through an 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 applications, $\mathbf{x}$ 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*. Being unable to directly sense chemical gradients because of its reduced dimensions, 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 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 quite 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 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 controllability/reachability results for this problem and provide a controller that, under the appropriate controllability conditions, 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.

Although the control of probability densities is still an incipient subject in the control literature, a substantial body of related work can be found in the literature of Markov Chain Monte Carlo (MCMC) methods [9]. These methods use a Markov chain to sample from a known (but usually hard to compute) distribution and then estimate integrals associated with that distribution. MCMC is largely used in statistical physics and in bayesian inference. In fact, our method can be regarded

as an instance of a dynamical/hybrid Markov Chain Monte Carlo method [10]. In particular, the *hit-and-run* method [11] resembles optimotaxis in that it also executes a piecewise linear random walk. The main difference between optimotaxis and traditional MCMC is that samples can be discarded in MCMC, which is not possible in our case due to the physical nature of the process.

This paper is organized as follows: the description of the problem is given in Section 2; Section 3 provides some auxiliary results on Markov processes that are useful in the control design and stability proofs; the proposed controller is described in Section 4; examples are given in Section 5; conclusions and final comments are given in Section 6.

II. PROBLEM DESCRIPTION

Initially, we briefly describe the concept of Piecewise-Deterministic Markov Processes (PDP) that is used in the paper, following closely the framework introduced in [12]. 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 $x \in \Omega := \mathbb{R}^d$ and $v \in V$, where V is a compact set. During flows, $\mathbf{x}(t)^1$ evolves according to the vector field $f(x, v)$, whereas $\mathbf{v}(t)$ remains constant and only changes with jumps. For a fixed $v \in V$, we denote by $\varphi_t^v x$ the continuous flow at time t defined by the vector field $f(\cdot, v)$ and starting at x at time 0. The conditional probability that a jump occurs between the time instants t and s, $0 < s < t$, given $\mathbf{x}(s)$ and $\mathbf{v}(s)$, is

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

where $\lambda(x, v)$ is called the *jump rate* at $(x, v) \in \Omega \times V$. At each jump, $\mathbf{v}$ assumes a new value given by the *jump pdf* $T_x(\cdot, \mathbf{v}^-)$. Thus, if a jump occurs at time t_k, then $T_{\mathbf{x}(t_k)}(v, \mathbf{v}^-(t_k))$ is the probability density of $\mathbf{v}(t_k)$ at v given $\mathbf{x}^-(t_k)$ and $\mathbf{v}^-(t_k)$, where the superscript minus indicates the left limits of the respective processes.

This PDP model is captured by several stochastic hybrid system models that appeared in the literature, including our stochastic hybrid models discussed in [13], or the hybrid models initially proposed in [14] by Hu, Lygeros and co-workers and further expanded in a series of subsequent papers [15]. Fig. 1 depicts a schematic representation of our PDP.

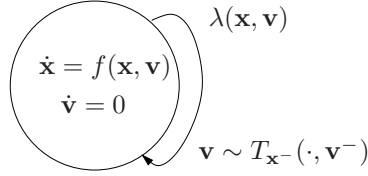[1]We use boldface symbols to indicate random variables.

Fig. 1. Hybrid automaton for the PDP

We define $p(x, v, t)$ as the joint probability density of the state (x, v) at time t. Here it is important to explicit the structure of the parameter space V. We consider V to be a compact subset of a locally compact separable metric space equipped with a Borel measure ν such that $\nu(V) = 1$. Note that, as opposed to [12], we do not require V to be countable. This more general setting for PDP's is supported by the theory developed in [11]. Denoting by m the Lebesgue measure in Ω, we have that $\int_{\Omega \times V} p(x, v, t) \, dm \times d\nu = 1, \ \forall t \geq 0$. We denote by $L^1(\Omega \times V)$ the space of real functions integrable with respect to $m \times \nu$.

In our setting, the vector field f is a given and the jump rate λ and the jump pdf T_x are control parameters. The controller cannot measure the state x directly; instead, an observation variable $\mathbf{y}(t) = g(\mathbf{x}(t))$ is given. In general, the function $g(x)$ is not known to the controller, which only has access to $y(t)$.

Assuming that $g(x)$ is nonnegative and integrable, our objective is to design λ and T such that a randomized controller will select $\mathbf{v}(t)$ as a function of the observations $\{\mathbf{y}(\tau); 0 \leq \tau \leq t\}$ collected up to time t so that the marginal $\int_V p(x, v, t) d\nu(v)$ converges to $cg(x)$, where c is a normalizing constant chosen so that $cg(x)$ integrates to one. As it will be clear later, it is not necessary for the controller to know the normalizing constant c in order to implement the proposed control law.

In practice, $g(x)$ is a chosen function of some physical measurements $F(x)$. 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

$$Q(F) = \begin{cases} F & \text{, if } F > \delta \\ \delta e^{F-\delta} & \text{, if } F \leq \delta \end{cases} , \tag{2}$$

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

$$Q(F) = F^n , \tag{3}$$

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 (2) above].

III. Some Fundamental Results in Stochastic Hybrid Systems

In this section we provide a few key results on the invariant probability densities of our PDP. The first result provides a generalized Fokker-Planck-Kolmogorov equation that governs the evolution of probability densities. We assume throughout the text that the vector field f is continuously differentiable on $\Omega \times V$, that there is no finite escape time, and that only a finite number of jumps occur in finite time intervals.

Theorem 1 ([16])**.** *If there exists a pdf $p(x, v, t)$ for $(\mathbf{x}(t), \mathbf{v}(t))$ that is continuously differentiable on $\Omega \times V \times \mathbb{R}^+$, then it satisfies the following generalized Fokker-Planck-Kolmogorov equation:*

$$\frac{\partial p}{\partial t} + \nabla_x \cdot fp = -\lambda p + \int_V T_x(v, v') \lambda(x, v') p(x, v', t) d\nu(v') , \tag{4}$$

where the divergence operator $\nabla_x \cdot$ is taken with respect to the variable x only. ☐

When $f(x, v) = v$, (4) has an important role in linear transport theory, where it models particles moving with constant velocity and colliding ellasticly [17], [18]. In this case, regarding p as the density of particles, (4) has a simple intuitive interpretation: on the left-hand side we find a drift term $\nabla_x \cdot vp$ 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, v), and an integral term corresponding to the particles jumping to the state (x, v).

Equation (4) 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 (4)] corresponds to an invariant marginal distribution $\int_V p(x, v, t) \ d\nu(v)$ that is proportional to $g(x)$. In fact, it will even be possible to obtain a *joint* invariant distribution $p(x, v, t)$ that is independent of v, therefore also proportional to $g(x)$. For simplicity of presentation, in the sequel we assume that $g(x)$ has been scaled so that $\int g(x)dx = 1$ and therefore $g(x)$ is already the desired invariant distribution. However, none of our results require this particular scaling.

A. Elements of the Ergodic Theory for Markov Chains

Once we derive control laws that result in the desired invariant density, it will be necessary to verify whether or not $p(x, v, t)$ actually converges to it from an arbitrary initial distribution. For this purpose, we present some results from ergodic theory that are useful to characterize the convergence of our PDP. Because ergodic theory is more thoroughly developed for discrete-time processes, we first state the results for a general discrete-time Markov chain and then show how to adapt them to continuous time. A second reason for discussing discrete-time processes is that, in practice, measurements are sampled at discrete time instants thus defining a discrete-time process.

Consider a general discrete-time Markov chain $\{\xi_k\}$ on the measurable space $(\mathcal{Y}, \mathcal{B})$, where $\mathcal{Y}$ is a locally compact separable metric space and $\mathcal{B}$ is the corresponding Borel σ-algebra. The chain is defined using the *transition probability function P*:

$$P(y, B) = \Pr\{\xi_{k+1} \in B | \xi_k = y\} \ , \tag{5}$$

for all $y \in \mathcal{Y}, B \in \mathcal{B}$ and $k \geq 0$. We say that μ is an *invariant probability measure* for P if, for every $B \in \mathcal{B}$,

$$\mu(B) = \int_{\mathcal{Y}} P(z, B)d\mu(z) \ . \tag{6}$$

A set $B \in \mathcal{B}$ is said to be *invariant with respect to P* if $P(x, B) = 1$ for all $x \in B$. An invariant probability measure is said to be *ergodic* if, for every invariant set B, $\mu(B) = 0$ or $\mu(B) = 1$. The following version of the law of large numbers for Markov chains will allow us to conclude that ergodicity of the invariant measure is sufficient to allow us to estimate $g(x)$ by sampling $\mathbf{x}(t)$ over time.

Theorem 2 ([19],Thm. 5.4.1). *Suppose that μ is ergodic. Then, there exists a measurable set $\mathcal{Y}' \subset \mathcal{Y}$ such that $\mu(\mathcal{Y}') = 1$ and, for every initial condition $y \in \mathcal{Y}'$,*

$$n^{-1} \sum_{k=0}^{n-1} \psi(\xi_k) \rightarrow \int_{\mathcal{Y}} \psi \; d\mu \quad a.s. \tag{7}$$

for every test function $\psi \in L^1(\mu)$. $\qquad\qquad\square$

By applying Theorem 2 to polynomial test functions ψ, one concludes that the time averages that appear in the left-hand side of (7) can be used to construct consistent estimators for the moments of the stationary measure of the process. Further, this result also provides a methodology to construct a consistent estimator for the invariant measure itself. To achieve this, we define the empirical measure $\mu^{(n)}$ by

$$\mu^{(n)}(B) = n^{-1} \sum_{k=0}^{n-1} 1_B(\xi_k) \tag{8}$$

for every $B \in \mathcal{B}$, where 1_B is the indicator function of the set B. Thus, since the left-hand side of (7) is equal to the expected value of ψ with respect to the empirical measure $\mu^{(n)}$, we have that (7), when restricted to the set of bounded continuous test functions, gives precisely the definition of weak convergence of $\mu^{(n)}$ to μ [19]. We formulate this result in the following corollary.

Corollary 1. *If μ is an ergodic measure, then the empirical measure $\mu^{(n)}$ converges weakly to μ almost surely for every initial condition in $\mathcal{Y}'$.* $\qquad\qquad\square$

By taking the expectation in Corollary 1 we have that also the law of the process converges weakly to μ in the Cesàro sense (i.e., convergence of the partial averages). In this paper convergence of measures will only be considered in the Cesàro sense. Even though some results in the paper can be proven for convergence in total variation, we have limited practical interest in this type of convergence so we will not prove it here.

For continuous-time processes $\{\theta_t\}$, one can analogously consider a transition probability $P_t(y, B) = \Pr\{\theta_t \in B | \theta_0 = y\}$ and then define invariant measures and invariant sets as those that are invariant under P_t for all $t \geq 0$. The definition of an ergodic measure remains the same as above. Given the existence of an ergodic measure, one can state ergodic theorems analogous to the one above provided that a continuity condition for P_t is satisfied [20]. For our PDP, in

particular, one concludes that, if we construct a control law under which $g(x)$ defines an ergodic measure μ for the PDP, then the probability measure $P_t(y, \cdot)$ for $(\mathbf{x}(t), \mathbf{v}(t))$ converges weakly to μ in the Cesàro sense for almost all initial condition y in the support of $g(x)$. Moreover, for every initial probability density $p(x, v, 0)$ absolutely continuous with respect to $g(x)$, the Cesàro average $t^{-1} \int_0^t p(x, v, s) \, ds$ converges to $g(x)$ in total variation.

In practice, more important than the convergence in continuous time is the convergence of the sampled processes. For the continuous-time process $\{\theta_t\}$, let $\{\xi_k\} := \{\theta_{\tau_k}\}$ be the process defined by sampling $\{\theta_t\}$ at a sequence of times $\{\tau_k\}$, where $\tau_k \to \infty$ a.s. as $k \to \infty$. We say that a sampled chain $\{\xi_k\}$ is *well-sampled* if $\{\theta_k\}$ and $\{\xi_k\}$ have the same ergodic measures. We can formulate the following corollary for well-sampled chains.

Corollary 2. *If μ is an ergodic measure for $\{\theta_t\}$, then the conclusions in Theorem 2 hold for every well-sampled chain $\{\xi_k\}$.*

Thus, we can use well-sampled chains to recover the steady-state statistics of continuous-time processes. Notice, however, that periodic sampling schemes do not always produce well-sampled chains. Suppose that μ is an ergodic measure for $\{\theta_t\}$. If $\tau_k = kT_1$, for some positive constant T_1, then P_{T_1} is the corresponding transition probability for $\{\xi_k\}$ and μ is an invariant measure for P_{T_1}, but μ is not necessarily ergodic. In fact, invariant sets for P_{T_1} may not be invariant for P_t for all $t > 0$. We say that the process is *aperiodic* if, for all $T_1 > 0$, the invariant measures for P_{T_1} are also invariant for P_t for all $t > 0$. Clearly, for aperiodic processes, the process sampled according to $\tau(k) = kT_1$ is well-sampled.

All the cases we deal with in this paper involve aperiodic processes. Indeed, we derive controllability conditions that in a way demand aperiodicity. Yet, if one has to sample from a periodic process, it is possible to generate well-sampled chains by sampling at random times. For example, a well-sampled chain would be obtained if the sample times are chosen such that, for $k \geq 0$, $\tau_{k+1} - \tau_k$ are i.i.d. random variables with finite mean and probability distribution absolutely continuous with respect to the Lebesgue measure (see [21, Lemma 2.2]).

B. Ergocity for the PDP

The next result provides a sufficient condition for the ergodicity of the invariant measures of our PDP. Here, we call *jump Markov chain* the chain defined by the jumps alone, that is,

our PDP when the vector field f is set to 0. We say that the jump Markov chain is *irreducible* for a fixed $x \in \Omega$ if, for every initial condition $v \in V$ and every A with $\nu(A) > 0$, there is a positive probability that A will be eventually reached from v. We say that a measurable function $\psi : \Omega \times V \to \mathbb{R}$ is *path-continuous* (*path-differentiable*) if $\psi(\varphi_t^v x, v)$ is a continuous (continuously differentiable) function of t for all (x, v).

Assumption 1.   *i. the jump Markov chain is irreducible $\forall x \in \Omega$.*

ii. f is continuously differentiable in x and continuous in v.

iii. $\lambda(x, v)$ is uniformly bounded on $\Omega \times V$ and, for any bounded and path-continuous ψ it holds that

$$t \mapsto \lambda(\varphi_t^v x, v) \int_V T_{\varphi_t^v x}(v', v) \psi(\varphi_t^v x, v') \ d\nu(v') \tag{9}$$

is continuous.

iv. $span\{f(x, v); v \in V\} = \Omega, \forall x \in \Omega$.

$\square$

Theorem 3. *Under Assumption 1, suppose that there exists an invariant probability measure for the PDP, then this measure is ergodic.* $\square$

We note that condition *iv* can be relaxed as done in Example 2. Under the conditions of the theorem, one can replace the weak convergence results for the law of the process in the previous subsection by the stronger notion of convergence in the total variation norm. For that purpose, one can use the Phillips expansion to show that the PDP is a T-process [22]. Then the restriction of the process to the ergodic classes is a Harris process and convergence in total variation follows [23].

C. Proof of Theorem 3

For a bounded measurable function $\psi : \Omega \times V \to \mathbb{R}$, we define the *transition operator* of our PDP to be

$$P_t \psi(x, v) = \mathbf{E}_{(x,v)} \psi(\mathbf{x}(t), \mathbf{v}(t)) \ , \tag{10}$$

where $\mathbf{E}_{(x,v)}$ denotes the expectation with respect to the initial condition (x, v) at time 0. Given an invariant measure μ for the PDP, a function $\psi \in L^1(\mu)$ is said to be *invariant with respect to μ* if $P_t \psi = \psi$ μ-a.e. for all $t > 0$.

A key result that we use to establish ergodicity of an invariant measure is the fact that μ is ergodic if and only if every function that is invariant with respect to μ is constant μ-a.e. [19, Lemma 5.3.2]. Lemma 1 below provides two key properties of the invariant functions of our PDP. Namely, it states that invariant functions must be invariant both under flows and jumps. We will see later that, under certain conditions, only constant functions can be invariant under flows and jumps for our PDP. This will show that all invariant functions are constant and therefore that μ is ergodic.

Lemma 1. *Let μ be an invariant probability measure for the PDP and let $\gamma \in L^1(\mu)$ be an invariant function with respect to μ . Then, under Assumption 1 (iii),*

$$\gamma(\varphi_t^v x, v) = \gamma(x, v) \quad \forall t > 0, \ \mu(dx, dv)\text{-a.e.} \tag{11}$$

and

$$\gamma(x, v) = \gamma(x, v') \quad \bar{\nu}(dx, dv, dv')\text{-a.e. }, \tag{12}$$

where $\bar{\nu}(dx, dv, dv') = \lambda(x, v)T_x(v, v')\nu(dv') \times \mu(dx, dv)$. □

The proof of Lemma 1 is presented in the appendix.

Proof of Theorem 3:

Let μ denote the invariant measure and let $\gamma \in L^1(\mu)$ be invariant under P_t. From Assumption 1 (*iii*) we have that Lemma 1 applies. Since by Assumption 1 (*i*) the jump Markov chain is irreducible, we have from Lemma 1 that γ does not depend on v μ-a.e. Without loss of generality, let γ be such that it does not depend on v on $\operatorname{supp}\mu$ and such that (11) holds on $\operatorname{supp}\mu$.

Given $(x_0, v_0) \in \operatorname{supp}\mu$, let $\{(x_0, v_i)\} \in \operatorname{supp}\mu$ be such that $\operatorname{span}\{f(x_0, v_i), \ i = 1, \ldots, d\} = \mathbb{R}^d$, which is possible due to Assumption 1 (*ii*) and (*iv*). For $\tau := (t_1, \ldots, t_d) \in \mathbb{R}^d$, let $y = \varphi_{t_d}^{v_d} \circ \cdots \circ \varphi_{t_1}^{v_1} x_0$. Then, from Assumption 1 (*ii*) we have that

$$\left.\frac{\partial y}{\partial \tau}\right|_{\tau=(0,\ldots,0)} = \left[\ f(x_0, v_1) \quad \ldots \quad f(x_0, v_d) \ \right] . \tag{13}$$

Since $\det\left(\frac{\partial y}{\partial \tau}(0, \ldots, 0)\right) \neq 0$, there exists a neighborhood $B_r(x_0)$ such that every $y \in B_r(x_0)$ can be written as $y = \varphi_{t_d}^{v_d} \circ \cdots \circ \varphi_{t_1}^{v_1} x_0$ for some $\tau \in \mathbb{R}^d$. Because μ is invariant, $y \in \operatorname{supp}\mu$ and (11) in Lemma 1 implies that

$$\gamma(\varphi_{t_d}^{v_d} \circ \cdots \circ \varphi_{t_1}^{v_1} x_0) = \gamma(x_0) \ , \tag{14}$$

where the dependence on v was omitted. From (14), we have that γ is constant in $B_r(x_0)$. Since $x_0 \in S$ is arbitrary, γ must be constant μ-a.e. Thus, as commented above, μ is ergodic by [19, Lemma 5.3.2]. ∎

IV. Control Design

In this section we provide a family of control laws that achieve our objective.

A. A Controllability Condition

We start with a controllability analysis that provides necessary and sufficient conditions under which a steady-state solution $p(x, v, t) = h(x, v)$, $\forall (x, v) \in \Omega \times V$ and $t > 0$, may be enforced. Naturally, we assume that $\int_{\Omega \times V} h \ dm \times d\nu = 1$.

Theorem 4. *Given a density $h(x, v) > 0$, $\forall (x, v) \in \Omega \times V$, with $\nabla_x \cdot fh \in L^1(\Omega \times V)$, there exists a jump intensity λ and a jump pdf T_x such that h is an invariant density for the PDP if and only if*

$$\int_V \nabla_x \cdot fh(x, v) \ d\nu(v) = 0, \ \forall x \in \Omega \ . \tag{15}$$

Moreover, when this condition is verified, the PDP has the desired invariant distribution h for a uniform jump distribution $T_x(v, v') = 1$, $\forall x, v, v'$ and a jump intensity

$$\lambda(x, v) = \frac{\alpha(x) - \nabla_x \cdot fh(x, v)}{h(x, v)}, \tag{16}$$

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

To prove Theorem 4, it is convenient to define the following integral operator:

$$K\psi = \int_V T_x(v, v')\psi(x, v') \ d\nu(v') \tag{17}$$

$\psi \in L^1(\Omega \times V)$. This operator allows us to rewrite the Fokker-Planck-Kolmogorov equation (4) in the following compact form:

$$\frac{\partial p}{\partial t} + \nabla_x \cdot fp = -\lambda p + K(\lambda p) \ . \tag{18}$$

Proof of Theorem 4: Substituting $p(x, v, t) = h(x, v)$ in (18) and rearranging the terms, we obtain

$$\lambda h = -\nabla_x \cdot fh + K(\lambda h) \ . \tag{19}$$

If we regard λh as an unknown, (19) can be seen as a Fredholm integral equation of the second kind for every fixed x [24].

To prove necessity, we note that, since T_x is a probability kernel, we have $\int_V T_x(v, v') \, d\nu(v) = 1$, $\forall x \in \Omega$, which is the same to say that $1 \in \mathcal{N}(I - K^*)$, $\forall x \in \Omega$, where K^* denotes the adjoint of K acting on $L^\infty(\nu)$ and $\mathcal{N}(\cdot)$ denotes the null space. From (19), we have that $\nabla_x \cdot fh \in \mathcal{R}(I - K)$, where $\mathcal{R}(\cdot)$ denotes the range of the operator. The necessity of (15) is then a consequence of the "orthogonality" of $\mathcal{R}(I - K)$ and $\mathcal{N}(I - K^*)$ [25].

To prove sufficiency, we select T_x so that $I - K$ is Fredholm [24] with null space spanned by the function $l(x, v) > 0$, i.e., solutions to (19) have the form

$$\lambda h = -(I - K)^\dagger \nabla_x \cdot fh + \alpha(x)l(x, v) \ , \tag{20}$$

where $\dagger$ is used to denote the generalized inverse and $\alpha(x) \in \mathbb{R}$ is a design parameter.

One such choice is $T_x(v, v') = 1$, for which $(I - K)\nabla_x \cdot fh = \nabla_x \cdot fh$ by the controllability condition (15) and therefore $(I - K)^\dagger \nabla_x \cdot fh = \nabla_x \cdot fh$. In this case, $l(x, v)$ is a constant, which leads to

$$\lambda h = \alpha(x) - \nabla_x \cdot fh \ . \tag{21}$$

Note that by choosing $\alpha(x) = \max_{v \in V} |\nabla_x \cdot fh|$ one can obtain a nonnegative λ such that $\lambda h \in L^1(\Omega \times V)$. Therefore, there exist λ and T_x such that h is an invariant density. ∎

Remark 1. It may happen that the λ given by (21) is not uniformly bounded, which might be an issue in proving stability of the invariant density. A sufficient condition (which is also a necessary condition under appropriate hypotheses) to have $\lambda(x, v) < 2M$, $\forall(x, v)$, for some finite constant M, is $|\nabla_x \cdot fh| \leq Mh$, $\forall(x, v)$.

B. Output Feedback Controller

Let us consider now the amount of information that is needed to implement the control law with the above λ. We can rewrite (16) as

$$\lambda = h^{-1}\alpha(x) - f \cdot \nabla_x \ln h - \nabla_x \cdot f \ . \tag{22}$$

To compute $\lambda(x, v)$, the controller needs to evaluate three terms.

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

$$f \cdot \nabla_x \ln h(\mathbf{x}(t), \mathbf{v}(t)) = \frac{d \ln h}{dt^+}(\mathbf{x}(t), \mathbf{v}(t)) \ , \tag{23}$$

  where '+' denotes the derivative from the right. Therefore, if we make $h = g$, the controller only needs to have access to the time derivative of the observed output y 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.

- To evaluate the term $h^{-1}\alpha(x)$, the controller needs to know a bound on $|g^{-1}\nabla_x \cdot fg|$ at every x. State feedback may then be dispensed when this bound can be expressed as a known function of the observed output function g. This bound can actually be estimated on the run as discussed in Example 1.

In summary, λ can be implemented using output feedback under the condition that $\nabla_x \cdot f$ is known and independent of x and $\max_{v \in V} |g^{-1}\nabla_x \cdot fg|$ is a known function of g.

As a combination of the above discussion and Corollary 2 and Theorems 2, 3 and 4, we can formulate the following theorem.

Theorem 5. *Suppose that*

$$\int_V f \ d\nu(v) = 0 \tag{24}$$

and $span\{f(\cdot, v); v \in V\} = \Omega$. *Then, for any* $g(x)$ *such that there exists a constant* $M < \infty$ *satisfying* $|\nabla_x \cdot fg| \leq Mg$, *the choice* $T_x(v, v') = 1$ *and*

$$\lambda(\mathbf{x}, \mathbf{v}) = M + \epsilon - \frac{d \ln \mathbf{y}}{dt^+} - \nabla_x \cdot f(\mathbf{x}, \mathbf{v}) \ , \tag{25}$$

for any $\epsilon > 0$, *implies that the probability measure defined by* $g(x)$ *is ergodic. Consequently, the averages* $t^{-1}\int_0^t p(x, v, s) \ ds$ *converges to* $g(x)$ *in total variation as* $t \to \infty$ *for all initial probability densities* $p(x, v, 0)$. *Moreover, for any* $\tau > 0$ *and any* ψ *such that* $\psi g \in L^1(\Omega \times V)$,

$$n^{-1} \sum_{k=0}^{n-1} \psi(\mathbf{x}(\tau k), \mathbf{v}(\tau k)) \to \int_{\Omega \times V} \psi(x, v) \ g(x)dm(x) \times d\nu(v) \quad a.s. \tag{26}$$

for almost all initial conditions. $\square$

Remark 2. The role of condition (24) is to guarantee that the controllability condition (15) holds for every g. Interestingly, this condition also ensures that the process is aperiodic. To understand why, note that, under condition (24), for a given vector field $f(x, v)$ there exists a convex combination of vector fields $\{f(x, v'); v' \in V\}$ that equals $-f(x, v)$. Roughly speaking, from an initial condition (x, v), there are trajectories that return to (x, v) in arbitrarily small time by selecting a proper combination of modes $\{f(x, v'); v' \in V\}$ that cancels $f(x, v)$. Hence, because $\lambda > 0$ and $T_x = 1$, trajectories starting in any set with nonempty interior return to that set with positive probability in arbitrarily small time, which implies aperiodicity.

Remark 3. Theorem 5 admits two straightforward generalizations. The first generalization is to have M to be a bounded function of g. The second generalization involves different choices of $T_x(\cdot, \cdot)$: the conclusions of the theorem still hold when $T_x > 0$ and the operator K satisfies $K1 = 1$, $Kf = 0$ and $K\nabla \cdot f = 0$.

Remark 4. Though this technique is quite successful for problems that satisfy the controlabillity condition (24), it is still an open problem to design controllers under limited controllability and information. For example, consider the case in which $\Omega = \mathbb{R}^2$, $V = [-1, 1]$, ν is a uniform probability measure, $y = g(x)$ and $f(x_1, x_2, v) = [x_2 \ v]^T$. It is straightforward to verify that the controllability condition (15) cannot be verified with $h(x, v) = g(x)$, $\forall(x, v)$. In this case, we need to have some v-dependent invariant density $h(x, v)$ such that $\int_V h(x, v) = g(x)$. However, it is not clear if that can be done using information from y only.

V. EXAMPLES

In this section we present applications of our main result to three systems caracterized 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 [27], which has turning constraints, but can reverse its direction of motion instantaneously. The third vehicle can be attracted/repelled by one of three beacons in the plane. Finally, we discuss how our results can be used to understand chemotaxis in the *E. coli*.

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 $x \in \Omega = \mathbb{R}^d$ and velocity $v \in V = \mathbb{S}^d$, the unit sphere. The measure ν is the Lebesgue measure on the sphere modulo a normalization factor. In this case we have $f = v$. Our objective is to make the probability density of the vehicles' position to converge to the observed function $g(x)$ and then have an external observer that can measure the vehicles position to collect information about $g(x)$. For this example Theorem 5 applies with

$$\lambda = \eta - v \cdot \nabla_x \ln g \ , \tag{27}$$

where $\eta > \|\nabla_x \ln g\|$. Since $\int_V \lambda \, d\nu = \eta$, one can interpret η as the average tumbling rate.

According to (1), the probability of a vehicle maintaining a run with the same direction in the interval $[0, t]$ is given by

$$\exp\left(-\int_0^t \lambda(\mathbf{x}(\tau), \mathbf{v}(\tau))d\tau\right) = \exp\left(-\int_0^t \eta - \frac{d}{d\tau}(\ln g(\mathbf{x}(\tau)))d\tau\right)$$
$$= e^{-\eta t} \frac{g(\mathbf{x}(t))}{g(\mathbf{x}(0))} \ . \tag{28}$$

This provides a simple and useful expression for the practical implementation of the algorithm: Suppose that an agent tumbled at time t_k. At that time pick a random variable ρ uniformly distributed in the interval $[0, 1]$ and tumble when the following condition holds

$$g(\mathbf{x}(t)) \leq \rho e^{\eta(t - t_k)} g(\mathbf{x}(t_k)), \ \ t \geq t_k \ . \tag{29}$$

As opposed to what (27) seems to imply, one does not need to take derivatives to implement (27). Also, the control law is not changed if a constant scaling factor is applied to $g(x)$, which is important because we could not be able to apply a normalizing constant to an unknown function g.

Another interesting feature is that η may be adjusted online. A vehicle may begin a search with $\eta = \epsilon > 0$ and if at some time t it observes that $\eta < \bar{\eta} := t^{-1} \ln g(\mathbf{x}(t))/g(\mathbf{x}(0))$, then it updates η to $\bar{\eta} + \epsilon$. The use of a small residue ϵ guarantees a positive λ. In this case, one can prove that the probability of the vehicle visiting any neighborhood in space is positive. Hence, η will eventually converge to $\sup \|\nabla \ln g\| + \epsilon$. A more elaborate adaptation can be obtained by

choosing η to depend on x through $g(x)$. With a space-dependent η, the conclusions in Theorem 5 would still hold and it would be possible to reduce the number of unnecessary velocity jumps.

We note that most physical quantities propagate with spatial decay not faster than exponential, which allows for the uniform boundedness of $\|\nabla_x \ln g\|$. 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 in Section 2) as long as a finite bound for their decay rate is known.

Ergodicity provides the basis for a procedure to estimate $g(x)$ by observing the position of N vehicles: We start by partitioning the region of interest into a family of sets $\{A_i \subset \Omega\}$, then we sample the vehicles' positions at times $k\tau \in \{0, \tau, 2\tau, \ldots, (M-1)\tau\}$, 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 $g(x)$ on the set A_i. To see why this is the case, we define

$$G_{N,M}(A_i) = \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{k=0}^{M-1} 1_{A_i}(\mathbf{x}_n(k\tau)) \ , \tag{30}$$

where x_n denotes the position of the n-th vehicle. Assuming that the agents have mutually independent motion, by Theorem 5 we have that

$$G_{N,M}(A_i) \rightarrow G(A_i) := \int_{A_i} g(x)dx \ \text{ a.s.} \tag{31}$$

as $M \rightarrow \infty$. This shows that $g(x)$ can be estimated by averaging the observations of the vehicles' position as in (30). The use of multiple agents ($N > 1$) improves the estimates according to the relation

$$\text{var}(G_{N,M}) = \frac{\text{var}(G_{1,M})}{N} \ . \tag{32}$$

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 main capability of optimotaxis, the localization of 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). When the equilibrium is reached, most agents concentrate in a neighborhood of the global

maximum. Yet, a portion of the agents clearly indicates the existence of the local maximum. We notice that the center of mass of the swarm goes straight through the local maximum to the global one. This feature is not shared with most deterministic optimization procedures and even with some stochastic ones. As a bonus, the information on secondary sources (local maxima) is not lost.
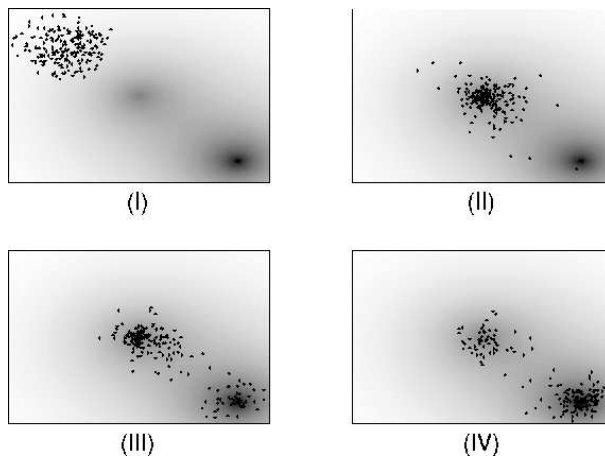

Fig. 2. Different stages of optimotaxis in the presence of two maxima. Black dots represent agents position whereas the background intensity represents the signal intensity. $F(x) = 0.4e^{-\|x\|} + 0.6e^{-\|x-[1.5 \ -1.5]'\|}$, $g(x) = F^n(x)$ with $n = 10$.

To quantify the convergence of the positions of the agents to desired distribution $g(x)$, we compute the correlation coefficient between the desired $G(A_i)$ and the empirical $G_{M,N}(A_i)$, when regarded as functions on $\{A_i\}$. 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. In the quantized case, we used the quantized version of the desired density $g(x)$ to calculate the coefficient. Interestingly, the addition of noise does not seem to affect considerably the transient response. Nevertheless, the residual error is greater due to the fact that the stationary density is not the one expected. On the other hand, quantization has a more negative impact on convergence time.

The sensitivity of the procedure with respect to the parameter n of the preprocessing function is studied with Fig. 4. The mean-square error of the vehicles position with respect to the maximum is used as a performance index. One notices that the performance degrades for n too low or too
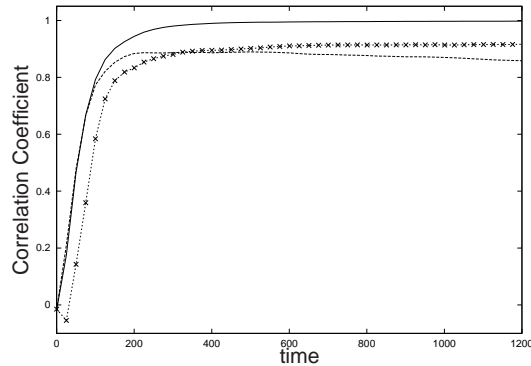
Fig. 3. Evolution of the coefficient of correlation for: the noiseless case (solid), the quantized measurements case (cross), and the exogenous noise case (dashed). The number of quantization levels is 64. The noise added to $\dot{v}$ is white Gaussian with standard deviation 10^{-2} along each axis. 100 agents were uniformly deployed in the rectangle $[-2.5, -1.5] \times [1.5, 2.5] \times V$. Refer to Fig. 2 for more details.

high. In particular, the sensitivity to noise and quantization increases with n. This suggests that an interesting strategy to reduce the effect of uncertainties and quantization is to assign agents with different values of n. In this case, the observed density would converge to an arithmetic average of the powers $F^n(x)$. Thus, the mean-square error would be smaller than the error corresponding to the maximum or minimum value of the chosen n.
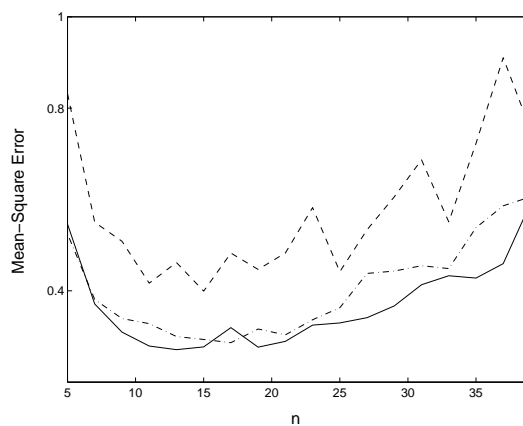

Fig. 4. Mean-square error with respect to the maximum of $F(x) = e^{-\|x\|}$ as a function of n. Noiseless case (solid), quantized $F(x)$ (dashed), and exogenous noise (dash-dotted). The number of quantization levels is 128. The noise added to $\dot{v}$ is white Gaussian with standard deviation 10^{-3} in each axis.

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

$$f(x, v) = \begin{bmatrix} v_1 \cos x_3 \\ v_1 \sin x_3 \\ v_2 \end{bmatrix}, \tag{33}$$

where $V = \{-v_0, 0, v_0\} \times \{-\omega_0, 0, \omega_0\}$ and ν is the uniform probability density over V. This kind of vehicle is referred to in the literature as the Reeds-Shepp car [27].

The vector field satisfies the controllability condition (24). Hence, we can use the same λ and T as in Theorem 5 to make g an invariant density. More precisely, $\lambda = \eta - f \cdot \nabla_x \ln g$ and $T_x = 1$. Note that, even though $\{f(x, v); \ v \in V\}$ no longer spans Ω, it is still easy to verify ergodicity using Lemma 1.

Given x and y in Ω, there is a trajectory linking these two points that consists of the vehicle spinning around (x_1, x_2) until it is aligned with (y_1, y_2), and then moving in a straight line to (y_1, y_2). Using Lemma 1 as in the proof of Theorem 3, one concludes ergodicity (alternatively, ergodicity is also implied by the fact that every invariant measure of an irreducible process is ergodic [19, Proposition 4.2.2]). Ergodicity would still hold true even if zero linear velocity was not allowed. For that case, note that $\varphi_t^v x$ defines a circular trajectory in Ω when $v_1 = v_0$ and $v_2 = \omega_0$. If (y_1, y_2) lies outside this circle, there exists a tangent line to the circle passing through (y_1, y_2). Thus, a trajectory from x to y consists of the vehicle moving along the circle and then along this tangent line until it reaches (y_1, y_2). If (y_1, y_2) lies inside the circle, then the vehicle only needs to move far enough from (y_1, y_2) before the procedure above can be executed. Finally, aperiodicity can be verified as in Remark 2.

Figure 5 illustrates how the empirical distribution indeed converges to the desired density. It shows that this 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 case for which 0 linear velocity is not allowed convergence is only slightly slower than the case in which it is allowed.
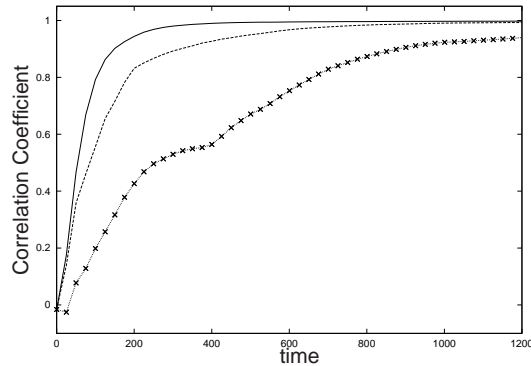
Fig. 5. Evolution of the coefficient of correlation for the unconstrained turning case (solid), for the constrained turning case with $v_0 = 1$ and $\omega_0 = 0.3$ (dots) and for the constrained turning case with $v_0 = 1$ and $\omega_0 = 0.15$ (cross).

C. Example 3

In this example vehicles make use of three beacons in order to navigate. In particular, vehicles always move straight towards or straight away from one of the beacons. Let $\Omega = \mathbb{R}^2$, $V = \{a, b, c\} \times \{-1, 1\}$, where a, b, c (the position of the beacons) are three points in $\mathbb{R}^2$ not in the same line, and ν is the uniform probability distribution over V. We take $f(x, v)$ to be $f = v_2(x - v_1)$. Thus, we have three points in the plane that may be either stable or unstable nodes. This is an example for which the divergence is not zero. According to Theorem 5, we choose

$$\lambda = \eta - f \cdot \nabla \ln g - 2v_2 \ , \tag{34}$$

for some η sufficiently large. Note that f satisfies the hypotheses of Theorem 5, since a, b and c are not aligned. The class of reachable densities includes those satisfying $\|\nabla \ln g\| \le M x^{-1}$, which includes all densities with polinomial decay. We note that a uniform T_x is not the only one that achieves the desired density for such a λ. For example, it is possible to choose T_x such that

$$T_x(v, v') = \frac{1}{4} 1_{V - \{v_1' \times \{-1, 1\}\}}(v) \ . \tag{35}$$

This jump pdf is such that jumps to the flows with the same fixed point are not allowed. Yet, since λT_x still defines an irreducible Markov chain, we can apply Theorem 3 to conclude uniqueness of the desired pdf and therefore convergence to it.

Remark 5. In [3] it is shown that the same type of objective of optimotaxis 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 5 would only be valid for vector fields that depend exponentially on the controlled parameters, which would make a diffusion controller solution to this last example very unlikely.

D. Chemotaxis

Chemotaxis in the bacterium *E. coli* is a good example of how the jump control of probability densities can be used for the optimal distribution of individuals. It is remarkable that the expression for λ in (27) obtained in the optimotaxis example is an affine function of $d(\ln \mathbf{y})/dt$. Hence, it coincides with simple biochemical models for the tumbling rate of the *E. coli*; see, for instance, Alt [2, Equation 4.8]. This author essentially proposed the existence of a chemical activator for the locomotion mechanism such that a tumble would occur each time the concentration of this activator would become less 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 (27). 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 (27)], which would modulate the run length. Though the use of tumble and run in optimotaxis is inspired by chemotaxis, one would not necessarily expect that our choice of the tumbling rate would lead to control laws that resemble the biochemical models in bacteria. 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 suppose 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. Given that $p(x, v, t)$ converges in total variation (see [3]), we have from [28] that

$$H(t) = - \int_\Omega \int_V p(x, v, t) \ln \left(\frac{1}{2} + \frac{1}{2} \frac{g(x)}{p(x, v, t)} \right) \, dx d\nu(v) \to 0 \;, \tag{36}$$

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 $H(t) \geq 0$ with equality 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 arises as an evolutionary equilibrium.

With our analysis, we hope to have shed some new lights regarding the functionalities of chemotaxis. Similar conclusions were drawn for predators in the work of [29]. However, we point out that further investigation is necessary given that our conclusions are based on a quite simplistic model for chemotaxis.

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 shown to be particularly useful when the observation process is a fixed but unknown function of the state. Controllability conditions were derived to determine when such a controller can enforce a given density to be a stationary density for the process. Specially, we were interested in the goal of making the probability density of the process converge to the observation function. We discussed potential applications of this theory in 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 to develop design tools for systems with limited controllability such as those with relative degree higher than or equal to one, as discussed in Remark 4. A second important problem is to define convergence rates in a manner that is useful for both analysis and design. A possible framework is provided by the theory of Large Deviations [30]. 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.

APPENDIX

PROOF OF LEMMA 1

For path-differentiable functions ψ we define the operator

$$\mathcal{L}_\varphi \psi(x, v) := \left.\frac{d}{dt} \psi(\varphi_t^v x, v)\right|_{t=0} . \tag{37}$$

In particular, if f is continuous and ψ is continuously differentiable, then $\mathcal{L}_\varphi \psi = f \cdot \nabla_x \psi$. We define a second operator on measurable functions $\psi : \Omega \times V \to \mathbb{R}$ by

$$\mathcal{K}\psi(x, v) = \lambda \int_V T_x(v', v)(\psi(x, v') - \psi(x, v)) \; d\nu(v') . \tag{38}$$

Under Assumption 1(*iii*), we have from [11, Chap. 7] that $A := \mathcal{L}_\varphi + \mathcal{K}$ is the *infinitesimal generator* of P_t, i.e.,

$$\lim_{t \to 0} \frac{P_t \psi - \psi}{t} = A\psi \tag{39}$$

for all ψ in the domain $D(A)$ of the operator, which is characterized by all bounded path-differentiable ψ such that $A\psi$ is bounded and path-continuous. To prove Lemma 1, we will extend P_t to act as a semigroup of operators on $L^1(\mu)$, where μ is the invariant probability measure in the lemma. Because $D(A)$ is dense in $L^1(\mu)$, P_t defines a strongly continuous semigroup on $L^1(\mu)$ with generator $\mathcal{A}$ such that $D(A) \subset D(\mathcal{A})$ and $A = \mathcal{A}$ on $D(A)$ [31]. Using [31, Thm. 2.4], one can proceed as in the proof of [11, Thm. 7.8.2] to show that

$$\int_{\Omega \times V} \mathcal{A}\psi \; d\mu = 0 \tag{40}$$

for all $\psi \in D(\mathcal{A})$.

Proof of Lemma 1:

A consequence of the dual ergodic theorem [19, Thm 2.3.6] is that invariant functions under P_t correspond to the Radon-Nikodym derivative of an invariant measure with respect to the invariant measure μ. Hence, we have that $\gamma\mu$ is also an invariant probability measure for P_t. Consequently, $\mu_1 := (\alpha_1 + \beta_1 \gamma)\mu$ and $\mu_2 := (\alpha_2 + \beta_2 \gamma)\mu$ are also invariant probability measures, where α_i and β_i are positive constants such that $\alpha_i + \beta_i = 1$, $i = 1, 2$, and $\alpha_1 \neq \alpha_2$. Define $\bar\gamma$ as the Radon-Nikodym derivative of μ_2 with respect to μ_1:

$$\bar\gamma := \frac{d\mu_2}{d\mu_1} = \frac{\alpha_2 + \beta_2 \gamma}{\alpha_1 + \beta_1 \gamma} . \tag{41}$$

By the dual ergodic theorem [19, Thm 2.3.6], $\bar{\gamma}$ is invariant under P_t, i.e., $P_t\bar{\gamma} = \bar{\gamma}$ μ-a.e. In addition, $0 < \min(\alpha_2/\alpha_1, \beta_2/\beta_1) \leq \bar{\gamma} \leq \max(\alpha_2/\alpha_1, \beta_2/\beta_1) < \infty$.

By [11, Lemma 7.7.3] and the invariance of $\bar{\gamma}$, we have that

$$\bar{\gamma}(x, v) = \bar{\gamma}(\varphi_t^v x, v) \exp\left(-\int_0^t \lambda(\varphi_s^v x, v) \, ds\right) + \int_0^t ds\, \lambda(\varphi_s^v x, v) \exp\left(-\int_0^s \lambda(\varphi_u^v x, v) \, du\right)$$
$$\cdot \int_V T_{\varphi_s^v x}(v', v)\bar{\gamma}(\varphi_s^{v'} x, v') \, d\nu(v') \quad \mu\text{-a.e.}$$
(42)

Taking the derivative in t, one can conclude that $\mathcal{A}\bar{\gamma} = A\bar{\gamma} = 0$ and, since $\bar{\gamma}$ is bounded, $\mathcal{L}_\varphi\bar{\gamma}$, $\mathcal{K}\bar{\gamma} \in L^1(\mu)$. From this and the uniform bound on $\ln\bar{\gamma}$, we conclude that $\mathcal{L}_\varphi \ln\bar{\gamma}$, $\mathcal{K}\ln\bar{\gamma} \in L^1(\mu)$ as well. Thus, $\ln\bar{\gamma} \in D(\mathcal{A})$ and $\mathcal{A}\ln\bar{\gamma} = A\ln\bar{\gamma}$. Consider the function

$$\mathcal{H} := \int_{\Omega \times V} A\ln\bar{\gamma} \, d\mu_2 - \int_{\Omega \times V} \mathcal{A}\bar{\gamma} \, d\mu_1 \; .$$
(43)

Using the definition of A, we can expand $\mathcal{H}$ as

$$\mathcal{H} = \int_{\Omega \times V} \left[\frac{\mathcal{L}_\varphi\bar{\gamma}}{\bar{\gamma}} + \lambda \int_V T_x(v', v)(\ln\bar{\gamma}(x, v') - \ln\bar{\gamma}(x, v)) \, d\nu(v') \right.$$
$$\left. - \bar{\gamma}^{-1}\mathcal{L}_\pi\bar{\gamma} - \lambda\bar{\gamma}^{-1} \int_V T_x(v', v)(\bar{\gamma}(x, v') - \bar{\gamma}(x, v)) \, d\nu(v') \right] d\mu_2$$
$$= \int_{\Omega \times V} \lambda \left[\int_V T_x(v', v) \left(1 + \ln\frac{\bar{\gamma}(x, v')}{\bar{\gamma}(x, v)} - \frac{\bar{\gamma}(x, v')}{\bar{\gamma}(x, v)} \right) d\nu(v') \right] d\mu_2 \; . \quad (44)$$

However, according to (40), $\mathcal{H} = 0$. Since $1 + \ln a - a \leq 0$ with equality only for $a = 1$, we conclude from (44) that $\bar{\gamma}(x, v) = \bar{\gamma}(x, v')$ $\bar{\nu}$-a.e. But

$$\gamma = -\frac{\alpha_2 - \alpha_1\bar{\gamma}}{\beta_2 - \beta_1\bar{\gamma}}$$
(45)

implies that also $\gamma(x, v) = \gamma(x, v')$ $\bar{\nu}$-a.e. This proves (12). From the definition of $\mathcal{K}$, we have $\mathcal{K}\bar{\gamma} = 0$ μ-a.e. Since $A\bar{\gamma} = 0$ μ-a.e., it follows that also $\mathcal{L}_\varphi\bar{\gamma} = 0$ μ-a.e. This proves (11) for $\bar{\gamma}$. Using (45), we can extend (11) to γ. ∎

REFERENCES

[1] H. Berg and D. Brown. Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. *Nature*, 239(5374):500–504, October 1972.

[2] W. Alt. Biased random walk models for chemotaxis and related diffusion approximations. *J Math Biol*, 9(2):147–177, April 1980.

[3] A. R. Mesquita, J. P. Hespanha, and K. J. Åström. Optimotaxis: A stochastic multi-agent optimization procedure with point measurements. In Magnus Egerstedt and Bud Mishra, editors, *Hybrid Systems: Computation and Control*, number 4981, pages 358–371. Springer-Verlag, Berlin, March 2008. Available at http://www.ece.ucsb.edu/~hespanha/published.

[4] D. A. Hoskins. Least action approach to collective behavior. In L. E. Parker, editor, *Proc. SPIE Vol. 2593, p. 108-120, Microrobotics and Micromechanical Systems, Lynne E. Parker; Ed.*, volume 2593 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 108–120, December 1995.

[5] M. Vergassola, E. Villermaux, and B.I. Shraiman. Infotaxis as a strategy for searching without gradients. *Nature*, 445(7126):406–409, 2007.

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

[7] A. Dhariwal, G.S. Sukhatme, and A.A. Requicha. Bacterium-inspired robots for environmental monitoring. In *IEEE International Conference on Robotics and Automation*, pages 1436–1443, New Orleans, Louisiana, Apr 2004. IEEE.

[8] A. Linhares. Synthesizing a predatory search strategy for VLSI layouts. *Evolutionary Computation, IEEE Transactions on*, 3(2):147–152, 1999.

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

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

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

[12] M.H.A. Davis. *Markov models and optimization*. Monographs on statistics and applied probability. Chapman & Hall, London, UK, 1993.

[13] J.P. Hespanha. Modeling and analysis of stochastic hybrid systems. *IEE Proc — Control Theory & Applications,* Special Issue on Hybrid Systems, 153(5):520–535, 2007.

[14] J. Hu, J. Lygeros, and S. Sastry. Towards a theory of stochastic hybrid systems. In N.A. Lynch and B.H. Krogh, editors, *Hybrid Systems: Computation and Control*, volume 1790 of *LNCS*, pages 160–173. Springer, 2000.

[15] ML Bujorianu and J. Lygeros. General stochastic hybrid systems: Modelling and optimal control. In *43rd IEEE Conference on Decision and Control, 2004. CDC*, volume 2, 2004.

[16] J. Bect. A unifying formulation of the fokker-planck-kolmogorov equation for general stochastic hybrid systems, 2008.

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

[18] M. Mokhtar-Kharroubi. *Mathematical Topics in Neutron Transport Theory*. World Scientific, Singapore, 1997.

[19] O. Hernandez-Lerma and J. B. Lassere. *Markov Chains and Invariant Probabilities*, volume 211 of *Progress in Mathematics*. Birkhauser, Basel, 2003.

[20] U. Krengel. *Ergodic Theorems*, volume 6 of *de Gruyer Studies in Mathematics*. Walter de Gruyter, Berlin; New York, 1985.

[21] P. Tuominen and R.L. Tweedie. The recurrence structure of general Markov processes. *Proc. London Math. Soc*, 39:554–576, 1979.

[22] S.P. Meyn and RL Tweedie. Stability of Markovian processes II: Continuous-time processes and sampled chains. *Advances in Applied Probability*, pages 487–517, 1993.

[23] S.P. Meyn and R.L. Tweedie. *Markov chains and stochastic stability*. Springer, 1996.

[24] R. Kress. *Linear integral equations*. Springer Verlag, 1999.

[25] N. Dunford and J. Schwartz. *Linear Operators*, volume VII of *Pure and Applied Mathematics*. Interscience Publishers, New York, 1957.

[26] K. Pichór and R. Rudnicki. Continuous markov semigroups and stability of transport equations. *Journal of Mathematical Analysis and Applications*, 249:668–685, 2000.

[27] H.J. 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*, 9110, 1991.

[28] J. Lin. Divergence measures based on shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, January 1991.

[29] P. Kareiva and G. Odell. Swarms of Predators Exhibit" Preytaxis" if Individual Predators Use Area-Restricted Search. *American Naturalist*, 130(2):233, 1987.

[30] J.D. Deuschel and D.W. Stroock. *Large Deviations*. American Mathematical Society, 2001.

[31] A. Pazy. *Semigroup of Linear Operators and Applications to Partial Differential Equations*. Springer-Verlag, New York, 1983.