UNIVERSITY OF CALIFORNIA
Santa Barbara

Exploiting Stochasticity in Multi-agent Systems

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical and Computer Engineering

by

Alexandre Rodrigues Mesquita

Committee in Charge:

Professor João P. Hespanha, Chair

Professor Mustafa H. Khammash

Professor Upamanyu Madhow

Professor Andrew R. Teel

December 2010

The Dissertation of
Alexandre Rodrigues Mesquita is approved:

Professor Mustafa H. Khammash

Professor Upamanyu Madhow

Professor Andrew R. Teel

Professor João P. Hespanha, Committee Chairperson

December 2010

Exploiting Stochasticity in Multi-agent Systems

to the engineer Antônio Mesquita.

Acknowledgements

To begin with, I would like to thank the members of the committee, who are the first readers of this text. I thank Profs. Madhow and Khammash for their advice on the early stages of this work. I thank Prof. Teel for his teaching on hybrid systems and for his kindness in being part of the committee despite of circumstances of my late invitation. I thank my advisor Prof. João Hespanha for bringing me to UCSB when I was supposed to go to a colder place. I also thank João for the opportunity to learn from his extraordinary thinking and working.

I am indebted to Prof. Karl Åström for his enthusiastic teaching and for his contribution to the early results in Chapter 3 and Chapter 4. Prof. Sean Meyn had an important influence in the ideas of Chapter 5. I also thank Prof. Girish Nair for his collaboration in the results of Chapter 6. Discussions with Prof. Vijay Gupta were useful in deriving the results of Section 6.3.

During these four years, I was positively motivated by the teaching of many professors including Roy Smith, Petar Kokotović, Charles Akemann, Mihai Putinar, Francesco Bullo, Igor Mezić. The CCDC seminars were a truly valuable resource and I would like to thank all the people who kept them running.

I thank all my lab mates and lab visitors for our fruitful (or not so much) interaction. In special, I would like to remember Payam, Naveen, Paolo, Abhi, Prabir, Duarte, Pietro, Rodrigo, Andrea, Jason, Josh, Steve, Farshad, Danny and Shaunak.

For excellent entertainment, I would like to thank my climbing partners, my fellow Brazilians and the St. Mark's people. For their patience, I thank my family.

Finally, I would like to acknowledge the financial and logistic support of the CAPES Foundation (Brazil) and of the Fulbright Commission. My research was also supported by the Inst. for Collaborative Biotechnologies through grants DAAD19-03-D-0004 and W911NF-09-D-0001 from the U.S. Army Research Office, and by the NSF grant CNS-0720842.

Curriculum Vitæ

Alexandre Rodrigues Mesquita

Education

2005 – 2006	MSc in Electronics Engineering and Computer Science, Instituto Tecnológico de Aeronáutica, São José dos Campos, Brazil
2000 – 2004	BSc in Electronics Engineering, Instituto Tecnológico de Aeronáutica, São José dos Campos, Brazil

Experience

2006 – 2010	Graduate Research Assistant, University of California, Santa Barbara.
2006	Internship for Casimiro Montenegro Filho Foundation, São José dos Campos, Brazil
2004	Young researcher fellowship, The State of São Paulo Research Foundation (FAPESP), São José dos Campos, Brazil
2003	Teaching of Brazilian and Portuguese Literature, CASD Vestibulares, São José dos Campos, Brazil

Selected Publications

Mesquita, A. R., Hespanha, J. P. and Nair, G. N.: "Redundant Data Transmission in Control/Estimation Over Lossy Networks", Submitted to Automatica.

Mesquita, A. R. and Hespanha, J. P.: "Jump Control of Probability Densities with Applications to Autonomous Vehicle Motion," *IEEE Trans. in Automatic Control*, (to appear).

Mesquita, A. R. and Hespanha, J. P.: "Construction of Lyapunov Functions for Piecewise-Deterministic Markov Processes," In *Proc. 49th Conf. on Dec. and Control*, December 2010.

Mesquita, A. R., Hespanha, J. P. and Nair, G. N.: "Redundant Data Transmission in Control/Estimation Over Wireless Networks" In *Proc. of the 2009 Amer. Contr. Conf.*, June, 2009.

Mesquita, A. R., Hespanha, J. P. and Åström, K. J.: "Optimotaxis: A Stochastic Multi-agent Optimization Procedure with Point Measurements," In *Proc. Hybrid Systems: Computation and Control*, March 2008.

Mesquita, A. R., Rempel, E. L. and Kienitz, K. H.: "Bifurcation analysis of attitude control systems with switching-constrained actuators," *Nonlinear Dynamics*, Vol. 51, 2008.

Mesquita, A. R., Kienitz, K. H. and Rempel, E. L.: "Robust Limit Cycle Control in an Attitude Control System with Switching-Constrained Actuators," In *Proc. of the 47th Conf. on Decision and Contr.*, December 2008.

Abstract

Exploiting Stochasticity in Multi-agent Systems

Alexandre Rodrigues Mesquita

To control multi-agent systems one can exploit the recurrence properties of stochastic processes. We illustrate this principle through two applications. In both applications, systems are modeled as hybrid systems where Markov transitions on the discrete variables depend on the continuous variables.

In the first application, stochasticity is introduced to overcome uncertainty about the environment. Inspired by bacterial chemotaxis, we design algorithms that control the spatial distribution of mobile agents that can take point measurements of a monitored function but cannot measure their own positions. Applications include source-seeking, monitoring and deployment. We prove that the probability density of agents is led to converge exponentially to a predetermined function of the measurements, much like in Markov Chain Monte Carlo methods. In the process of designing these control algorithms, we prove results on piecewise deterministic Markov processes that can find application outside this particular design problem.

In the second application, we control the level of stochasticity in a networked control system. Given that the probability of successful communication can be significantly increased by transmitting multiple copies of the same message, we show that, by dynamically assigning the number of transmitted copies of the same data, one can obtain significant performance gains with only a modest increase in the total number of transmissions. We develop techniques to design communication protocols that exploit the transmission of multiple packets, while seeking a balance between stability/estimation performance and communication rate. An average cost optimality criterion is employed to obtain a number of optimal protocols applicable to networks with different computational capabilities. Other capacity scheduling techniques are also explored.

Contents

List of Figures

Chapter 1

Introduction

In the recent years, teams of autonomous vehicles have been considered to solve a number of problems including environmental monitoring, agricultural tasks such as crop spraying, search and rescue in disaster response, law enforcement and surveillance. While the technology for the construction of unmanned vehicles is now relatively mature, endowing these machines with the intelligence necessary to realize the above tasks is an extremely hard problem, which has become one of the major challenges in the fields of control and artificial intelligence. The reason for such complexity lies in the fact that agents have access only to local information or to information that is exchanged using faulty communication networks. This imperfect nature of information typically makes the computation of optimal solutions intractable. In this monograph, we propose solutions to two distributed control problems that have the potential to reduce the impact of these issues in multi-agent applications.

In our first problem, uncertainty in a spatial environment (that agents can only sense locally) is overcome by the introduction of stochasticity. In the same way that randomized policies are often needed in adversarial games, we propose that agents use randomized solutions to "play" against an uncertain environment. This is also the principle utilized in randomized optimization algorithms, which introduce stochasticity to avoid solutions that become trapped by local optima. Using only local information, we are able to induce agents into an ergodic behavior

in such a way that the frequency agents visit a certain region is proportional to the "importance" of that region to the task being executed.

The practical impossibility of computing optimal solutions for many multi-agent systems has motivated researchers to look for solutions in nature, taking into consideration that living organisms have spent millions of years specializing in the realization of similarly complex tasks. Following these lines, our first solution takes inspiration from the bacterium *E. Coli*, one of the most ancient living organisms.

Our second problem deals with communication failures, and we show that agents may benefit by treating the capacity of communication channels as a variable to be controlled. Essentially, agents are able to control the statistics of the communication channel depending on the criticality of the information being transmitted. This solution is more suited to multi-agent systems rather than to traditional networked control systems because in this domain missions are typically heterogenous in time and the importance of the data being transmitted may vary significantly with time. Moreover, due to the agents mobility, channel quality may also change and, therefore, agents can profit greatly from adapting the use of communication resources (power, bandwidth, etc.) to realize energy savings and improve the lifetime of their energy sources.

Our solution to this problem draws inspiration from the adaptive techniques in the 3G and 4G networks. In these standards, significant performance gains are obtained by adapting transmissions to the state of the channels using techniques such as multiple antennas (MIMO), dynamic channel allocation and channel dependent scheduling. In our solution, however, we are adapting transmissions both to the stochastic realization of the channel and to the stochastic realization of the process being controlled.

To motivate our solutions, consider Figure 1.1, where a team of vehicles is tasked to find the source of a radio signal. Agents have limited sensing in that they can only infer their distance to the source using the strength of the measured signal. Instead of requiring all vehicles to converge to the source, we are satisfied

if we achieve a high density of vehicles in a neighborhood of the source. Our first solution allows agents to accomplish this objective independently, without direct cooperation, and is suited for cases in which agents are unable to measure their own position or to communicate with each other.

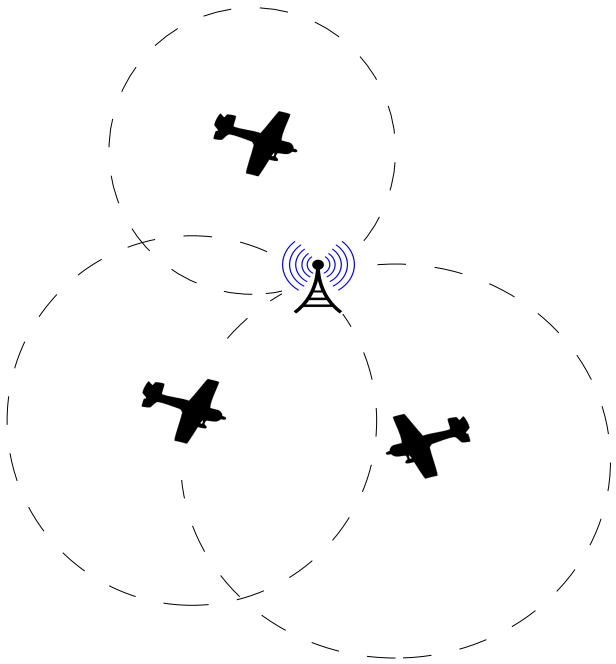

Figure 1.1: Source-seeking by a team of agents using range measurements

Suppose now that agents are able to exchange their position information and distances to the source. In this case, they can infer their relative position with respect to the source by calculating the intersection point of the dashed circles. However, failures in communication may render agents into a temporary "open loop" mode. Our second solution proposes that agents can modulate their ability to successfully transmit messages among themselves (e.g., by simultaneously using several independent communication channels) to avoid or remedy sequential communication failures and therefore keep the distance to the target bounded.

1.1 Dissertation Overview

Next, we provide a brief summary of the contributions in this mongraph.

1.1.1 Multi-agent Solutions by Controlling Agents Probability Density

We consider agents whose only observation is a scalar non-negative output $\mathbf{z} = q(\mathbf{x})$, where $\mathbf{x}$ is the state of the agent and $q(\cdot)$ is some function. Both $\mathbf{x}$ and $q(\cdot)$ are unknown to the agent. The control objective is to achieve a steady-state probability density for the state $\mathbf{x}$ that matches the function $Z(q(\cdot))$ up to a normalization factor, where Z is a function chosen by the designer. The state $\mathbf{x}$ typically includes the position of the agent, which can take point measurements $\mathbf{z} = q(\mathbf{x})$ at its current location.

This type of control objective can be used to solve numerous problems in the area of mobile robotics under very limited measurement models. Indeed, our formulation does not require agents to communicate or to have measurements of their current position or of the gradient of q.

In *deployment problems*, a group of such agents 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. A natural choice here is $Z(q)$ to be a monotone function that maps the measurement value to the desired density.

In *search problems*, a group of agents 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 $q(x)$ 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. The natural design choice for $Z(q)$ is some power q^n, where the exponent $n > 1$ should be large to accentuate the global maxima of q.

In *monitoring problems*, 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. A possible design choice is $Z(q) = q$.

1.1.2 Jump Control of Probability Densities

To accomplish the above control objective, we must design some control $\mathbf{m}$ such that the state $\mathbf{x}$, governed by the differential equation

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{m}(t)) \ ,$$

has its probability density controlled to the desired function. One of the novelties of our approach is the use of a stochastic supervisor that decides the probability that one should switch between different control values m, which is inspired by bacterial chemotaxis.

Suppose that the control space is some compact set $\mathbb{M}$ equipped with a probability measure ν such that $\operatorname{supp} \nu = \mathbb{M}$. Then, we show that a necessary and sufficient condition for the existence of a solution to the control problem is that 1) $\int_{\mathbb{M}} f(x, m) \ \nu(dm) = 0$ for all x and 2) the control system $\dot{x} = f(x, m)$ is approximately controllable, i.e., one can steer the state from any initial condition to any given open set.

Under these conditions, we provide a simple algorithm that is proven to accomplish our control objective. For the case in which f has zero divergence, this algorithm reads as follows.

1. Initialize the algorithm with $\boldsymbol{\tau}_0 = t = 0$, $k = 0$ and some constant $\eta > 0$.

2. Draw a random variable $\bar{\mathbf{m}}$ from the distribution ν and set $\mathbf{m}(t) = \bar{\mathbf{m}}$.

3. Draw a random variable $\mathbf{r_k}$ uniformly distributed in the interval $[0, 1]$.

4. When the output $\mathbf{z}(t)$ satisfies

$$Z(\mathbf{z}(t)) \leq \mathbf{r}_k \; e^{\eta(t-\boldsymbol{\tau}_k)} Z(\mathbf{z}(\boldsymbol{\tau}_k)), \;\; t \geq \boldsymbol{\tau}_k \;\; , \tag{1.1}$$

increment k, set $\boldsymbol{\tau}_k = t$ and return to step 2.

In proving the convergence of such algorithms, we derive novel results that contribute to the ergodic theory of Piecewise-Deterministic Markov Processes. We present a new method for the construction of Lyapunov functions for such processes based on the maximization of a certain notion of rate of convergence. This method allowed us to construct a Lyapunov function to prove exponential ergodicity of the controlled process.

1.1.3 Bio-inspiration and New Insights on Chemotaxis

The construction of our stochastic supervisory control draws inspiration from the chemotactic behavior of the bacterium *E. coli*, i.e., the mechanism by which this bacterium responds to chemical stimuli in order to move toward higher concentrations of nutrients. In turn, our results allow one to advance new conclusions about bacterial chemotaxis.

The fact that bacteria use a control similar to (1.1) indicates that bacteria evolved towards the control of probability densities. For bacteria, this objective may be useful to either find food as in search applications or, as in deployment applications, to distribute the population of bacteria so to maximize the amount of nutrients per individual.

Our results also support the conclusion that controls like (1.1) are intimately related to the experimental observation that bacterial chemotaxis obeys Weber's law, which states that the absolute response of a sensory system to a stimulus is proportional to the logarithm of the stimulus.

As a practical application, we employ our results to predict the behavior of recently proposed bacterial actuators, where bacteria are used to move a microstructure attached to their cell surface.

1.1.4 Communication Protocols Using Redundant Transmissions

To prove the potential of adaptive techniques in networked control systems (NCS), we consider the problem of controlling a linear time-invariant plant over a network of erasure channels. As depicted in Figure 1.2, the sensor has the freedom to transmit redundant copies of the same message through a number of independent channels. Our main concern is the design of protocols that select the number of redundant channels as a function of the system history. As our main contribution, we show that channel adaptive techniques, which are already used in data networks, can provide significant performance gains in control networks.

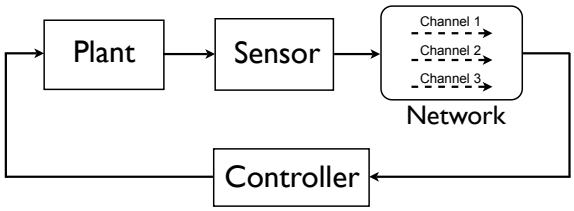

Figure 1.2: Architecture of the Networked Control System

The basic intuition behind this technique is that one can use redundancy to increase the probability of a successful transmission whenever the estimation error in the controller becomes large. On the other hand, at time instants for which the control performance is satisfactory, one may send only one packet or not send data at all, which would save communication resources. This adaptive behavior is desirable for NCSs because it improves the reliability of transmissions without relying on error correction schemes that induce delay in the transmissions. Indeed,

if a packet containing some measurement data is dropped at a time instant, it is generally more important from a control point of view to guarantee that the measurements at the next time instant are delivered, rather than to retransmit the old information that was dropped previously.

If M is the number of available channels and p is the drop out probability for each independent channel, we show the existence of protocols that stabilize the state covariance if

$$a^2 p^M < 1 \ , \tag{1.2}$$

where a is the spectral radius of the system matrix. Moreover, no stabilizing protocol exists if

$$a^2 p^M > 1 \ .$$

This implies that any process can be stabilized given a large enough number M of redundant channels. Under mild assumptions, we also show that the *same exact* stability criterion is valid for an arbitrarily large number of plants sharing the redundant channels.

To design protocols that take into account the conflicting objectives of minimizing the estimation error covariance in the controller and minimizing the communication resources expended by the sensor, we adopt an average cost criterion that minimizes a weighted cost of error covariance and communication rate. We show that optimal protocols exist if (1.2) is satisfied and that these protocols can be well approximated by protocols that use quadratic value functions and that can be constructed using numerically efficient procedures.

To implement the optimal protocols above, a sensor must compute the estimation error made by the controller. Since this may be an issue depending on the size of the state of the process or on the computational capabilities of sensors, we also propose optimal protocols that are allowed to depend only on the number of consecutive transmission failures. As numerical results show, both type of protocols can improve control performance with little additional communication costs.

The proposed technique has a diminishing returns property in the sense that little additional benefits are obtained by increasing the number of redundant channels beyond two or three. This is so because the probability of failure p^M becomes indistinguishably small for M larger than some small constant, typically 3. This implies that, as opposed to what one might expect, implementing a large number of channels is not necessary to obtain significant performance gains using our technique.

1.1.5 A Framework for General Capacity Scheduling

Beyond redundant data transmissions, we propose a general framework to exploit the idea of capacity scheduling. Consider the control system

$$\begin{aligned} \mathbf{x}(k+1) &= A\mathbf{x}(k) + B\mathbf{u}(k) + \mathbf{w}(k) \\ \mathbf{y}(k) &= C\mathbf{x}(k) + \mathbf{b}(k)\boldsymbol{\omega}(k) + \mathbf{v}(k) \ , \end{aligned}$$

where $\mathbf{x}$ is the state, $\mathbf{y}$ is the output received by the controller, $\mathbf{u}$ and $\mathbf{b}$ are control inputs, $\mathbf{w}$ is the process noise, $\mathbf{v}$ is the measurement noise and $\boldsymbol{\omega}$ is the noise resulting from the network transmission and A, B and C are the system matrices. In this formulation, we can regard the probability density of $\mathbf{b}(k)\boldsymbol{\omega}(k)$ as being controlled by means of varying the transmit power, the quantization or coding scheme, or the number of repeated packets. The cost associated with using a certain control $\mathbf{b}$ will depend on the method used to control capacity.

Since solving for optimal controls for this type of system is computationally hard, we exploit the linearity of the system to propose computationally tractable suboptimal policies.

1.2 Literature Review

In this section we provide a brief review of methods in multi-agent systems with a focus on the similarities and differences with respect to our approaches.

1.2.1 Markov Chain Monte Carlo and alike

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 [38]. 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 [72], our jump control approach can be regarded as a dynamical/hybrid MCMC type method. In particular, the *hit-and-run* method [90] resembles some of our algorithms 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.

An interesting distinction is given by the classic Metropolis-Hastings algorithm [38]. This algorithm samples the space using a random walk and generates a process with distribution $Z(q)$ by discarding samples according to a given rule. Remarkably, the rule to discard samples is exactly the same as (1.1). The distinction is that our algorithm cannot physically discard samples; instead, it "discards" the control input.

The well-known optimization method of simulated annealing is also commonly cast as a MCMC technique in which the target distribution q^n is progressively changed by increasing n to infinity along iterations [88]. A similar shaping of the target distribution is adopted here for search problems.

Our work also has connections with the field of reinforcement learning, especially with TD or Q-learning where unknown value functions are identified using only local observations of the cost function [15], 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 [61] and subsequent works.

1.2.2 Deployment and Environmental Monitoring

An overview of the problems and solutions in the fields of distributed robotics and swarm robotics can be found in [18] and [78], respectively. The main reasons to utilize this type of solutions are the robustness to failures in individual agents and the ability to operate simultaneously in different regions of the space, which makes the execution of more complex tasks possible.

Our deployment problem corresponds to the coverage control problem described in [18]. Typical solutions involve partitioning the space in Voronoi cells, placing agents at their center and updating the partition in order to optimize some coverage criterion. The main drawback of this type of technique is its communication and computational cost and the use of position information.

This coverage problem of mobile sensor networks has been considered in many other papers, most notably in [46], where a potential field is used to control the agents distribution. In [37], extremum seeking was considered to solve a deployment problem in which one desires to distribute agents on the neighborhood of the peak of the measured field.

In monitoring applications, mobile sensors replace a network of fixed sensors in situations that would demand an extremely high number of spatial samples. The main focus here is on deciding how to sample the space, in a problem analogous to the sensor deployment in fixed networks. A simple gridding technique was utilized to sweep the space in [85], where the problem of pollution monitoring using a robot equipped with an electronic nose was tackled experimentally. However, uniform grids are typically inefficient and more sophisticated sampling methods have been proposed in a series of papers dealing with the so-called *Networked Info Mechanical System (NIMS)*, a cable-based robotic system used to monitor forests [11, 75, 80, 81]. To improve upon gridding methods, adaptive sampling schemes

are proposed in [75] and [81], where one does not need to have information on the environment beforehand. A path planning problem with the goal of maximizing the collected information is solved in [80]. A heuristic sampling method motivated by bacterial chemotaxis was proposed in [28].

Our solution to environmental monitoring is distinct in that it is based on MCMC. As such, the function being monitored is sampled more often at points with higher value (or importance). This is especially interesting if one must calculate integrals of the monitored function. Indeed, MCMC is regarded as the computational method of choice to integrate a function in high dimensions [38].

1.2.3 Source-Seeking and Chemotaxis

Classical techniques from numerical optimization have been adapted for single- and multi-vehicle search strategies when gradients are not explicitly available [19, 82, 63]. In [19], the local gradient is estimated by means of a circular movement. The simplex method is implemented with a network of autonomous vehicles in [82]. However, this approach requires the ability to measure the vehicles' relative position. Mayhew et al. [63] proposed a hybrid control algorithm to perform a conjugate directions method without position measurements. This control was further extended to non-holonomic vehicles in [64].

Many other approaches have also been considered. A sliding mode control is presented in [62]. A deterministic approach for ascending/descending potential fields is proposed in [10] with control expressions that are reminiscent of ours. Control algorithms for networks of vehicles inspired by collective behavior such as fish schooling and chemotaxis are designed in [45, 9, 96]. Statistical approaches have also been proposed for the case of turbulent environments, but assuming the availability of vehicle's position measurements [74, 91].

An extremum seeking strategy is adopted in [98]. This strategy is reminiscent of our solution in that it utilizes a circular motion to indirectly estimate gradients.

In general, when convergence is proven in the above references, this is done exclusively under the assumption that the signal spatial profile is convex or quadratic-like.

Bio-inspired techniques have a strong appeal in optimization. Examples are the well-known genetic algorithms and the solutions for the traveling salesman problem inspired by ant colonies [32]. Inspiration by chemotaxis is not a new approach to the source-seeking problem, see e.g. [45, 91, 71, 28]. In [45], chemotaxis inspires the use of an Iterated Function Systems approach to lead a group of agents to equilibrium. Agents are also able to communicate in this work. In [91], agents move in the direction of the maximum expected information gain. In [71], a neural network approach is used to program nematode-like robots to perform chemotaxis. In [28], a biased random walk inspired by *E. coli*'s chemotaxis is implemented in robots.

Whereas the previous works on chemotaxis-based source-seeking rely on a heuristic approach and do not give convergence guarantees, our technique allows one to effectively control the probability density of the vehicle's position to a specified density whose peak coincides with the maximum of the measured signal.

1.2.4 Transmission Scheduling

Several authors (e.g. [7, 22, 47, 93, 4]) have explored the idea of saving communication resources in a NCS by scheduling transmissions in a judicious way. However, these works do not consider the possibility of redundant transmissions or other means of scheduling capacity.

In [7], the concept of Lebesgue sampling is introduced as an alternative to periodic (Riemann) sampling. The problem of minimizing the estimation error given a constraint on the total number of transmissions is considered in [47]. Most notably, [93] formulates transmission scheduling as a Markov Decision Process approach as in this monograph. A similar approach is considered in [22], where a suboptimal policy that approximates the optimal one by a constant factor is

derived. With respect to [93, 92], we obtain theoretical advances that prove the existence of optimal policies even for the case of packet drop outs. In [4], actuation times are scheduled as opposed to transmission times.

When scheduling transmissions, one should be aware that the very action of not transmitting data already conveys information to the controller, which raises an issue of whether the separation principle holds or whether a Kalman filter solution is optimal. The optimality of a Kalman-like filter was proven for a scalar process in [60] using symmetry arguments.

1.2.5 Adaptive Techniques in Data Networks

Many modern communication systems use channel feedback in order to improve performance. At the channel level, adaptation is typically achieved by adjusting the transmit power, by adaptive coding and by *diversity* schemes, which consist of the transmission of redundant signals through mostly independent channel realizations. Diversity schemes may involve using multiple time slots, frequency slots, antennas or network paths [86, 39].

Adaptation is especially advantageous in mobile networks to cope with time-varying fading in channels. Such adaptive techniques require that an estimate of the channel be made at the receiver and then sent to the transmitter. Based on the channel state, the transmitter allocates communication resources. Two basic adaptive techniques present in GSM and CDMA systems as well as wireless LANs are adaptive modulation and coding [39]. Channel feedback is also particularly important in MIMO systems, where multiple transmit and receive antennas are used.

In multiuser systems, techniques referred to as multiuser diversity or channel dependent scheduling assign the right of transmission to the users with the best channel gains in order to maximize throughput.

The diversity techniques above exploit the availability of multiple propagation paths with diverse quality. Analogously, networks often present multiple paths

with diverse performance. The use of path diversity for media streaming was first proposed in [5]. A scheduling problem for streaming media using path diversity was addressed in [20]. This use of diversity is similar to ours in the sense that it is application aware.

A number of adaptive techniques for data networks have used a Markov decision process framework similar to ours. In [21], an optimal rate-distortion problem for media streaming is considered. The problem of transmission scheduling over fading channels was considered in [53], where the cost was a tradeoff among average transmission power, average delay and average packet dropping. The problem of optimal power and rate allocation in MIMO systems over Markovian fading channels was considered in [30].

The main distinction of our work with respect to the above is that we do adaptation for control applications, which have different requirements than data transmission or media streaming, instability being one of the main concerns.

Chapter 2

Stochastic Hybrid Systems: Stability and Control

Stochastic Hybrid Systems (SHS) are traditionally thought of as Markov processes evolving on a space that includes both continuous and discrete variables. In this monograph, we consider SHSs in a broader sense, namely, these are Markov processes whose state space is the product of spaces with inherently different topologies. The case of continuous and discrete variables is certainly an instance of this definition. A second instance that we deal with in this monograph are hybrid systems in which some variables are defined on the Euclidian space and others are defined on the unit sphere. We address these systems under the framework of general state-space Markov processes. This chapter describes the basic concepts of this theory, along with the main results concerning stability and control of these processes. A class of hybrid system that we will pay particular attention to is that consisting of Piecewise-Deterministic Markov Processes.

2.1 Markov Processes in General Spaces

We consider a time-homogeneous Markov process $\mathbf{\Phi}(t)$[1] taking values in a locally compact separable metric space Y equipped with a Borel σ-algebra $\mathcal{B}$. Because most of the concepts and results are analogous for discrete and continuous time,

[1] In this monograph we will consistently use boldface symbols to denote random variables.

we present them simultaneously and hence assume that $t \in \mathbb{T}$, where $\mathbb{T}$ may be either $\mathbb{Z}^+$ or $\mathbb{R}^+$. For every $t \in \mathbb{T}$, we assume that $\mathbf{\Phi}(t)$ is a random variable on the measurable space $(\Omega, \mathcal{F})$, where $\mathcal{F} = \sigma\{\mathbf{\Phi}(t);\ t \geq 0\}$, and, for every $y \in \mathsf{Y}$, there is a probability measure $\mathbb{P}_y$ on $(\Omega, \mathcal{F})$ such that $\mathbb{P}_y\{\mathbf{\Phi}(0) \in A\} = 1_A(y)$, the indicator function of $A \in \mathcal{B}$. Expectations with respect to $\mathbb{P}_y$ are denoted by E_y. The process is described using the *transition semigroup* $\{P^t\}_{t\in\mathbb{T}}$, where P^t are the transition kernels

$$P^t(y, A) := \mathbb{P}_y\{\mathbf{\Phi}(t) \in A\}, \quad y \in \mathsf{Y}, A \in \mathcal{B} \ ,$$

and P^t are assumed to be measurable functions of y for fixed t and A. Viewed as an operator, the kernel P^t operates on measurable functions $h : \mathsf{Y} \to \mathbb{R}$ and signed measures μ on $\mathcal{B}$ according to

$$P^t h(y) := \int_{\mathsf{Y}} h(z) P^t(y, dz) \ \text{ and } \ \mu P^t(A) := \int_{\mathsf{Y}} P^t(y, A)\mu(dy) \ .$$

Acting either on functions or measures, these operators form a semigroup that satisfies the Chapman-Kolmogorov identity $P^{t+s} = P^t P^s$. A more intuitive interpretation is given by the equalities

$$P^t h(y) = \mathrm{E}_y\{h(\mathbf{\Phi}(t))\},$$

and

$$\mu P^t(A) = \mathrm{Pr}_\mu\{\mathbf{\Phi}(t) \in A\},$$

where Pr_μ denotes the probability given that $\mathbf{\Phi}(0)$ has distribution μ. We also assume that $\mathbf{\Phi}(t)$ is measurable on $\Omega \times \mathbb{T}$, from which it follows that $t \mapsto P^t(y, A)$ is measurable for fixed y and A.

To address long term properties of the process we define the *resolvent chain* $\boldsymbol{\xi}(n) := \mathbf{\Phi}(\boldsymbol{\tau}_n)$ as the process obtained by sampling $\{\mathbf{\Phi}(t)\}$ at a sequence of times $\{\boldsymbol{\tau}_n\}$, where $\boldsymbol{\tau}_{n+1} - \boldsymbol{\tau}_n$ are i.i.d. random variables with probability measure a_θ, for some parameter $\theta > 0$. For discrete time, $a_\theta(\tau) := (1-\theta)e^{-\theta\tau}$, and, for continuous time, $a_\theta(d\tau) = \theta e^{-\theta\tau}\ell(d\tau)$, where ℓ denotes the Lebesgue measure. The resulting

resolvent chain has a one-step transition probability given by the *resolvent kernel*

$$R_\theta = \int_{\mathbb{T}} P^t \; a_\theta(dt) \;\; .$$

Intuitively, the resolvent kernel can be regarded as a Laplace transform of the transition kernel. As a discrete-time process, the resolvent chain is an instrument that allows the extension of many discrete-time results to the continuous-time setting.

2.1.1 Invariant Measures and Ergodic Theory

Two of our main concerns in this monograph are the steady-state behavior of the processes and the consistency of measurements obtained from the process. In this section we define the concept of invariant measures and discuss how they relate with the time averages of measurements of a process.

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

$$\mu = \mu P^t \;\; \forall t \geq 0 \;\; .$$

A set $A \in \mathcal{B}$ is said to be *invariant with respect to* P^t if $P^t(y, A) = 1$ for all $y \in A$ and $t > 0$. An invariant probability measure μ is said to be *ergodic* if, for every invariant set A, $\mu(A) = 0$ or $\mu(A) = 1$. In particular, it is true that every invariant measure is a convex combination of ergodic measures [42, Chapter 5].

The following theorem is a version of the law of large numbers for discrete-time Markov processes. Its fundamental importance lies in the fact that observations of a particular realization of the process are consistent with its ergodic measures in the sense that time averages of observables converge almost surely to the ensemble average.

Theorem 1 ([42], Thm. 5.4.1)**.** *Suppose that μ is an ergodic probability measure for $\mathbf{\Phi}(n)$. Then, for every test function $h \in L^1(\mu)$, there exists a measurable set $\mathsf{Y}_h \subset \mathsf{Y}$ such that $\mu(\mathsf{Y}_h) = 1$ and, for every initial condition $y \in \mathsf{Y}_h$,*

$$n^{-1} \sum_{k=0}^{n-1} h(\mathbf{\Phi}(k)) \to \int_{\mathsf{Y}} h \; d\mu \quad \mathbb{P}_y\text{-a.s.} \tag{2.1}$$

Moreover, for the set of bounded continuous test functions, the result holds with some set Y' *independent of* h*.*

By applying Theorem 1 to polynomial test functions h, one concludes that the time averages that appear in the left-hand side of (2.2) 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

$$\mu^{(n)}(A) = n^{-1} \sum_{k=0}^{n-1} 1_A(\mathbf{\Phi}(k))$$

for every $A \in \mathcal{B}$. Thus, since the left-hand side of (2.2) is equal to the expected value of h with respect to the empirical measure $\mu^{(n)}$, we have that (2.2), when restricted to the set of bounded continuous test functions, gives precisely the definition of weak convergence of $\mu^{(n)}$ to μ [42]. 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* Y'*.*

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). This result is traditionally known as the mean-ergodic theorem.

For continuous-time processes, one can state ergodic theorems analogous to the one above [83, Thm. 6.1.26]. In practice, more important than the convergence in continuous time is the convergence of the sampled processes, which, for the aperiodic processes defined in the next section, always follows from the continuous-time convergence.

2.2 Drift Conditions for Ergodicity

In this section we recapitulate the concepts of irreducibility and ergodicity of a process and present some Foster-Lyapunov criteria to ensure ergodicity.

2.2.1 Irreducibility, Recurrence and Continuity

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

$$\psi(A) > 0 \Rightarrow R_1(y, A) > 0, \quad \forall y \in \mathsf{Y} \ .$$

We shall assume that ψ is a maximal irreducibility measure, i.e, any other irreducibility measure is absolutely continuous with respect to ψ. This notion of irreducibility is a generalization of irreducibility for stochastic matrices. In the context of stochastic matrices, ψ has positive measure for every element in the space.

For a ψ-irreducible process, we denote by $\mathcal{B}^+$ the set of measurable functions $s : \mathsf{Y} \to \mathbb{R}^+$ such that $\psi(s) := \int_{\mathsf{Y}} s(y)\psi(dy) > 0$. A function $s \in \mathcal{B}^+$ and a positive measure ν on $\mathcal{B}$ are called *small* if

$$R_\theta(y, A) \geq s \times \nu(y, A) := s(y)\nu(A), \quad y \in \mathsf{Y}, A \in \mathcal{B} \ ,$$

for some $\theta > 0$. We call $C \in \mathcal{B}$ a *petite* set if there is a small function s of the form $s = b1_C$, for a constant $b > 0$. Petite sets are numerous. Indeed, every set A such that $\psi(A) > 0$ contains a petite set [69]. In many applications, it turns out that every compact set is petite. This is the case, for example, for the irreducible T-processes defined later in this section.

We say that a ψ-irreducible process $\mathbf{\Phi}$ is *aperiodic* if, for some petite set C such that $\psi(C) > 0$, there exists a a time T such that $P^t(y, C) > 0$, for all $t \geq T$ and $y \in C$.

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

$$\boldsymbol{\eta}_A := \int_{\mathbb{T}} 1_A\{\mathbf{\Phi}(t)\}\ dt \ ,$$

where dt denotes the integration with respect to the Lebesgue measure or the counting measure for the continuous-time process and the discrete-time one, respectively. A ψ-irreducible process is called *recurrent* if $\mathrm{E}_y[\boldsymbol{\eta}_A] = \infty$ for all $y \in \mathsf{Y}$ and A such that $\psi(A) > 0$. The process is called *Harris recurrent* if $\mathbb{P}_y\{\boldsymbol{\eta}_A = \infty\} = 1$ for all $y \in \mathsf{Y}$ and A such that $\psi(A) > 0$. When $\boldsymbol{\Phi}$ is ψ-irreducible and admits an invariant probability measure, we say that $\boldsymbol{\Phi}$ is *positive*.

For positive Harris recurrent processes, we can establish stronger convergence results than those in Theorem 1.

Theorem 2 ([69], Thm. 17.0.1)**.** *Suppose that $\boldsymbol{\Phi}$ is a positive Harris recurrent chain with invariant probability measure μ. Then, for every test function $h \in L^1(\mu)$ and every initial condition y, we have that*

$$n^{-1}\sum_{k=0}^{n-1} h(\boldsymbol{\Phi}(k)) \to \int_{\mathsf{Y}} h \; d\mu \quad \mathbb{P}_y\text{-a.s.} \tag{2.2}$$

We say that $\boldsymbol{\Phi}$ is a *T-chain (process)* if, for some sampling distribution θ,

$$R_\theta(y, A) \geq T(y, A), \; y \in \mathsf{Y}, A \in \mathcal{B} \;\; ,$$

where $T(\cdot, A)$ is a lower semicontinuous function for all $A \in \mathcal{B}$ and $T(y, \mathsf{Y}) > 0$ for all $y \in \mathsf{Y}$. We say that $\boldsymbol{\Phi}$ is (weak) *Feller* if P^t maps the set of bounded continuous functions on Y into itself.

Through the following series of results, these two continuity properties play an important role in the analysis done in this monograph.

Proposition 1 ([69], Thm 6.0.1)**.** *If $\boldsymbol{\Phi}$ is a ψ-irreducible Feller chain such that* $\operatorname{supp}\psi$ *has non-empty interior, then $\boldsymbol{\Phi}$ is a ψ-irreducible T-chain.*

Proposition 2 ([69], Thm. 9.3.2)**.** *If $\boldsymbol{\Phi}$ is a T-chain (process) and $T(y, A) > 0$ for some $y \in \mathsf{Y}$, then there is a neighborhood O of y and a distribution θ such that*

$$\inf_{y' \in O} R_\theta(y', A) > 0$$

and, moreover, A is reached from O in finite time with positive probability.

A set H is called maximal absorbing if

$$y \in H \iff \{\boldsymbol{\eta}_H = \infty\} \ \ \mathbb{P}_y\text{-a.s.}$$

A set H is called *maximal Harris set* if H is maximal absorbing and $\boldsymbol{\Phi}$ restricted to H is Harris recurrent. One important property of T-processes is that they admit the following decomposition, commonly known as the Doeblin decomposition.

Theorem 3 ([67, 69], Thm. 3.4, Thm. 9.2.2)**.** *Suppose that $\boldsymbol{\Phi}$ is an irreducible T-process. Then, the state space may be decomposed into the disjoint union:*

$$\mathsf{Y} = H \cup E \ , \tag{2.3}$$

where H is a maximal Harris set with invariant measure μ and

$$\mathbb{P}_y\{\{\boldsymbol{\Phi} \to \infty\} \cup \{\boldsymbol{\eta}_H = \infty\}\} = 1, \ y \in \mathsf{Y} \ .$$

We prove next that E must be open.

Lemma 1. *The set E in the decomposition (2.3) 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 Proposition 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. □

2.2.2 W-Ergodicity and Exponential Ergodicity

Consider a function $V : \mathsf{Y} \to [1, \infty]$. For a measure μ on $\mathcal{B}$ and a function $\varphi : \mathsf{Y} \to \mathbb{R}$, we define the weighted V-norms

$$\|\mu\|_V = \sup_{|h| \leq V} \left| \int_{\mathsf{Y}} h \ d\mu \right| \quad \text{and} \quad \|\varphi\|_V = \sup_y \frac{|\varphi(y)|}{V(y)} \ .$$

For a real function $W \geq 1$ on Y, we say that $\mathbf{\Phi}(t)$ is *W-ergodic* if $\mathbf{\Phi}$ is positive Harris recurrent with invariant probability measure μ such that

$$\lim_{t\to\infty} \|P^t(y,\cdot) - \mu\|_W = 0 \quad \forall y \in \mathsf{Y} \ . \tag{2.4}$$

We say that $\mathbf{\Phi}(t)$ is *V-exponentially ergodic* if there exists $V : \mathsf{Y} \to [1,\infty]$, finite for at least one $y \in \mathsf{Y}$, an invariant probability measure μ, and constants $B_0, b_0 > 0$ such that

$$\|P^t(y,\cdot) - \mu\|_V \leq B_0 V(y) e^{-b_0 t}, \ \forall y \in \mathsf{Y}, t \in \mathbb{T} \ . \tag{2.5}$$

If $\mathbf{\Phi}$ is V-exponentially ergodic with $V(y)$ finite for all $y \in \mathsf{Y}$, then we say that $\mathbf{\Phi}$ is *exponentially ergodic*. If we denote by L_V^∞ the space of functions with bounded V-norm, we can regard P^t as a bounded linear operator in L_V^∞. We can then rewrite (2.5) as the convergence in operator norm:

$$\|P^t - 1 \times \mu\|_V \leq B_0 e^{-b_0 t}, \ t \in \mathbb{T} \ ,$$

where $\|\cdot\|_V$ denotes the induced norm in L_V^∞.

2.2.3 Drift Conditions for Ergodicity

We define the *extended generator* $\mathcal{L}$ of $\mathbf{\Phi}$ as follows. The domain $D(\mathcal{L})$ is defined as the set of functions $h : \mathsf{Y} \to \mathbb{R}$ such that there exists a measurable function $g : \mathsf{Y} \to \mathbb{R}$ for which the process

$$h(\mathbf{\Phi}(t)) - h(\mathbf{\Phi}(\mathbf{0})) - \int_0^t g(\mathbf{\Phi}(s)) \, ds$$

is a local Martingale with respect to the filtration $\{\mathcal{F}_t = \sigma(\mathbf{\Phi}(\mathbf{s}) : 0 \leq s \leq t)\}_{t\in\mathbb{T}}$. The value of the generator is then denoted $\mathcal{L}h := g$. For discrete time, the extended generator is simply $\mathcal{L} = P - I$, where I denotes the identity operator. For continuous time, this is an extension of the *infinitesimal generator* $\hat{\mathcal{L}}$, which is defined as

$$\hat{\mathcal{L}}h := \lim_{t\to 0} \frac{P^t h - h}{t}, \ h \in D(\hat{\mathcal{L}}) \subset B(\mathsf{Y}) \ ,$$

where $B(\mathsf{Y})$ is the set of bounded measurable Borel functions on Y equipped with the supremum norm.

We define the *drift conditions*:

(D1) For a small function s and a function $W \geq 1$, the function $V : \mathsf{Y} \to \mathbb{R}^+$ verifies

$$\mathcal{L}V \leq -W + s$$

and

(D2) For a constant $c > 0$, a small function s, the function $V : \mathsf{Y} \to [1, \infty)$ verifies

$$\mathcal{L}V \leq -cV + s \ .$$

The drift conditions allow us to write the following ergodic theorems.

Theorem 4 ([68, 55])**.** *Suppose* $\mathbf{\Phi}$ *is* ψ*-irreducible and aperiodic. Then, the following conditions are equivalent*

i. $\mathbf{\Phi}$ *is positive Harris recurrent with invariant probability measure* μ *and* $\mu(W) < \infty$*;*

ii. condition (D1) is satisfied for some function W*;*

iii. $\mathbf{\Phi}$ *is* W*-ergodic.*

Theorem 5 ([33, 55])**.** *Suppose* $\mathbf{\Phi}$ *is* ψ*-irreducible and aperiodic. Then, the condition (D2) satisfied for some function* V *is equivalent to* V*-exponential ergodicity.*

2.3 Piecewise-Deterministic Markov Processes

In this monograph we consider an important class of continuous-time Stochastic Hybrid Systems known as Piecewise-Deterministic Markov Processes (PDP). Our

definition of PDPs follows closely the framework introduced in [25] and extended in [48]. 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. The PDPs considered in this monograph are a simplified version of those defined in [25], since we do not allow processes to have deterministic jumps.

We consider state variables $x \in \mathsf{X} = \mathbb{R}^d$ and $m \in \mathbb{M}$ so that $\mathsf{Y} = \mathsf{X} \times \mathbb{M}$, where $\mathbb{M}$ is a compact set. During flows, the *continuous state* $\mathbf{x}(t)$ evolves according to the vector field $f(x, m)$, whereas the *discrete state* $\mathbf{m}(t)$ remains constant and changes only with jumps. For a fixed $m \in \mathbb{M}$, we denote by $\varphi_t(x, m)$ the continuous flow at time t defined by the vector field $f(\cdot, m)$ and starting at x at time 0. The conditional probability that at least one jump occurs between the time instants t and s, $0 < s < t$, given $\mathbf{x}(s)$ and $\mathbf{m}(s)$, is

$$1 - \exp\left(- \int_s^t \lambda(\varphi_{\tau - s}(\mathbf{x}(s), \mathbf{m}(s)), \mathbf{m}(s)) d\tau \right) \quad , \tag{2.6}$$

where the nonnegative function $\lambda(x, m)$ is called the *jump rate* at $(x, m) \in \mathsf{X} \times \mathbb{M}$. At each jump, the overall state $\boldsymbol{\xi} := (\mathbf{x}, \mathbf{m})$ assumes a new value distributed according to the *jump kernel* Q. Namely, if we define the process $\{\mathbf{T}_n\}$ to denote the time of the n-th jump, then

$$\Pr(\boldsymbol{\xi}(\mathbf{T}_k) \in A \mid \boldsymbol{\xi}^-(\mathbf{T}_k) = \xi) = Q(\xi, A)$$

for $A \in \mathcal{B}$, where the superscript minus indicates the left limit of a processes. With some abuse of notation, we use the symbol Q to denote the operator given by $Qh(y) = \int h(\xi) Q(y, d\xi)$ for $h \in B(\mathsf{Y})$.

This PDP model is captured by several stochastic hybrid system models that appeared in the literature, including the stochastic hybrid models discussed in [44], or the general stochastic hybrid models introduced in [17]. One main difference of PDPs with respect to general stochastic hybrid systems is that they do not allow

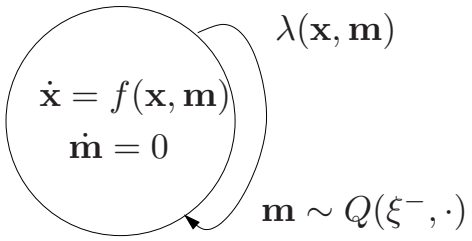

Figure 2.1: Hybrid automaton for the PDP

a diffusion term, i.e., they do not have a Wiener process component. Fig. 2.1 depicts a schematic representation of our PDP.

Under the standard assumptions in Assumption 1, it was proven in [25] that the defined PDP is a strong Markov process and a Borel right process [see [79] for the definitions].

Assumption 1. *i.* $x \mapsto f(x, m)$ *is locally Lipschitz for each* $m \in \mathbb{M}$ *and* φ_t *admits no finite escape time.*

ii. $\lambda : \mathsf{Y} \to \mathbb{R}^+$ *is a measurable function such that the map* $t \mapsto \lambda(\varphi_t(x, m), m)$ *is locally integrable.*

iii. The probability kernel Q *is such that* $Q(\xi, \{\xi\}) = 0$ *for* $\xi \in \mathsf{Y}$.

iv. $\mathrm{E}_\xi \left[\sum_{k=0}^{\infty} 1\{\mathbf{T}_k \leq t\}\right] < \infty$ *for all* $t > 0$ *and* $\xi \in \mathsf{Y}$.

Throughout this monograph, we assume that Assumption 1 holds for every PDP. Assumption 1 *(ii)* holds immediately when λ is uniformly bounded, which is true in most of the cases considered here. Since we do not consider the case of spontaneous jumps, it is always possible to redefine the jump rate λ so that Assumption 1 *(iii)* is satisfied. Assumption 1 *(iv)* also follows from the uniform boundedness of λ. Even when λ is unbounded but continuous, the possibility of infinitely many jumps in finite intervals can be ruled out by showing that the process is nonexplosive by considering stopped processes as in [68].

The extended generator of the PDP was first characterized in [25]. For path-differentiable functions h, define the operator

$$\mathcal{D}_\varphi h(x,m) := \frac{d}{dt} h(\varphi_t(x,m),m)\bigg|_{t=0} .$$

Then we can state the following theorem.

Theorem 6 ([25],§26.14)**.** *The domain of the extended generator of the PDP $\boldsymbol{\xi}$ consists of measurable functions h satisfying*

1. *The map $t \mapsto h(\varphi_t(x,m),m)$ is absolutely continuous for all $(x,m) \in \mathsf{Y}$.*
2. *The process*
$$\sum_{\mathbf{T}_n<t} h(\boldsymbol{\xi}(\mathbf{T}_n)) - h(\boldsymbol{\xi}^-(\mathbf{T}_n))$$
is locally integrable.

The extended generator is given by

$$\mathcal{L}h = \mathcal{D}_\varphi h + \lambda Q h - \lambda h \tag{2.7}$$

for $h \in D(\mathcal{L})$.

The second condition characterizing h on $D(\mathcal{L})$ is verified if, for example, $Qh - h$ is bounded. In the case h is differentiable, $\mathcal{D}_\varphi$ reduces to the more familiar $f \cdot \nabla_x$.

Next we discuss the adjoint of the generator, which defines a generalized Fokker-Planck-Kolmogorov equation that governs the evolution of probability densities. Here it is important to make explicit the structure of the space $\mathbb{M}$. We consider $\mathbb{M}$ to be a compact subset of a locally compact separable metric space equipped with a Borel probability measure ν such that $\operatorname{supp}\nu = \mathbb{M}$. With this structure, a function $p(x,m,t)$ satisfying $\int_{\mathsf{X}\times\mathbb{M}} p(x,m,t)\ell(dx)\nu(dm) = 1$ is a joint probability density of the state (x,m) at time t. Note that, as opposed to [25], we do not require $\mathbb{M}$ to be countable. This more general setting for PDPs is supported by the theory developed in [48].

To formulate the next result we need to strengthen Assumption 1.

Assumption 2. *It is assumed throughout this monograph that:*

i. f and $\nabla_x f$ are continuous functions on $\mathsf{X} \times \mathbb{M}$;

ii. There exists a kernel Q^* such that

$$\ell \times \nu(d\xi) Q(\xi, d\xi') = \ell \times \nu(d\xi') Q^*(\xi', d\xi) \ .$$

□

Theorem 7. *A continuously differentiable probability density $p(x, m, t)$ is a pdf for $(\mathbf{x}(t), \mathbf{m}(t))$ if and only if it satisfies the following generalized Fokker-Planck-Kolmogorov equation:*

$$\frac{\partial p}{\partial t} + \nabla_x \cdot fp = -\lambda p + Q^*(\lambda p) \ , \tag{2.8}$$

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

A derivation of this equation may be found in [36, Sec. 3.4]. A more general treatment for stochastic hybrid systems is given in [12].

When $f(x, m) = v_m \in [-1, 1]$, (2.8) has an important role in linear transport theory, where it models particles moving with constant velocity v_m and colliding ellasticly [52, 70]. In this case, regarding p as the density of particles, (2.8) has a simple intuitive interpretation: on the left-hand side we find a drift term $\nabla_x \cdot v_m p$ 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 kernel Q corresponding to the particles jumping to the state (x, m).

2.4 Markov Decision Processes

In this section we consider a discrete-time Markov control model $(\mathsf{X}, \mathsf{B}, \{B(x) : x \in \mathsf{X}, P, c\})$. The state space X and the *action space* B are both Borel sets; for a fixed x, $B(x) \subset \mathsf{B}$ denotes the *set of admissible actions* for the state x. We

assume that the set $\mathsf{K} := \{(x, b) : x \in \mathsf{X}, b \in B(x)\}$ is a Borel subset of $\mathsf{X} \times \mathsf{B}$. Given $b \in B(x)$, $P(\cdot|x, b)$ is a probability transition kernel on the state space X. The *one-step cost* c is a positive measurable real function on K.

A *control policy* π is a sequence of rules to choose admissible controls $\mathbf{b}(n)$ as a function of the process history $\{\mathbf{x}(n), (\mathbf{x}(k), \mathbf{b}(k)); 0 \leq k < n\}$. In general, control policies may be time-varying and stochastic. We denote by Π the set of all policies and by Π_0 the set of *stationary policies*, i.e., policies that satisfy $\mathbf{b}(n) = g(\mathbf{x}(n))$ for some measurable function $g : \mathsf{X} \to \mathsf{B}$ that verifies the constraint $g(x) \in B(x)$. For stationary policies, we write simply $\pi = g$.

For a fixed policy $\pi \in \Pi$ and an initial state $x \in \mathsf{X}$, there exist a stochastic process $(\mathbf{x}(n), \mathbf{b}(n))$ on K and a corresponding a probability measure $\mathbb{P}_x^\pi$ defined on the sample space $(\Omega, \mathcal{F})$, where $\Omega = (\mathsf{X} \times \mathsf{B})^\infty$ and $\mathcal{F}$ is the product σ-algebra, and $\mathbb{P}_x^\pi$ is compatible with the transition kernel P. We denote by E_x^π the corresponding expectation operator. In the case π is stationary, $\mathbf{x}(n)$ is a Markov process and we denote by $P_\pi(\cdot|x) := P(\cdot|x, \pi(x))$ the corresponding probability transition kernel. We also write $c_\pi(x) = c(x, \pi(x))$.

We define the following average cost (AC) minimization criterion

$$J(\pi, x) := \limsup_{N \to \infty} \frac{1}{N} \mathrm{E}_x^\pi \Big[\sum_{k=0}^{N-1} c(\mathbf{x}(k), \mathbf{b}(k)) \Big] \ .$$

A policy π^* is said to be *AC-optimal* if

$$J(\pi^*, x) = \inf_{\pi \in \Pi} J(\pi, x) =: J^*(x), \quad \forall x \in \mathsf{X},$$

and J^* is called the *optimal AC-function*. The goal of the theory of Markov Decision Processes (MDP) is to find optimal policies. In this monograph we consider mainly average cost criteria, but *discounted cost* criteria are also widely adopted in the MDP theory. The discounted cost criterion is given by

$$J_\gamma(\pi, x) := \mathrm{E}_x^\pi \Big[\sum_{k=0}^{\infty} \gamma^n c(\mathbf{x}(k), \mathbf{b}(k)) \Big] \ ,$$

where $\gamma < 1$ is a constant. Discounted cost problems and average cost problems are typically related in that, under the proper regularity conditions, the optimal policies for the discounted problems converge to the AC-optimal policy as $\gamma \to 1$.

The existence of AC-optimal policies is the subject of a vast literature. Here we present an approach that requires a somewhat minimal set of hypotheses.

Assumption 3. *There exists a measurable function $V : \mathsf{X} \to [1, \infty)$, a nontrivial positive measure ν on $\mathcal{B}(\mathsf{X})$, a nonnegative measurable function s on K, and a positive constant $\delta < 1$ such that*

i. $\|c\|_V < \infty$.

ii. $\nu(V) < \infty$.

iii. $P(C|x, b) \geq \nu(C)s(x, b), \quad (x, b) \in \mathsf{K}$ *and* $C \in \mathcal{B}(\mathsf{X})$.

iv. $PV \leq \delta V + s\nu(V)$.

v. $\nu(s(x, \pi(x))) > 0$ *for all* $\pi \in \Pi_0$.

Assumption 3 can be viewed as a drift condition (D2) that is satisfied uniformly on the set of policies. The following theorem establishes the ergodicity of the P_π chains.

Theorem 8 ([89])**.** *Suppose that Assumption 3 holds. Then, for $\pi \in \Pi_0$, $\mathbf{x}(n)$ is a ν-irreducible positive Harris chain with unique invariant measure μ_π. Moreover, $J(\pi, x) = \mu_\pi(c_\pi)$ and there exists a unique $\phi_\pi \in L_\infty^V$ satisfying $\nu(\phi_\pi) = 0$ and the Poisson equation:*

$$\mu_\pi(c_\pi) + \phi_\pi = c_\pi + P_\pi \phi_\pi \ .$$

The functions ϕ_π are known as relative value functions and they satisfy

$$\phi_\pi = \sum_{n=0}^{\infty} [P_\pi^n c_\pi - \mu_\pi(c_\pi)] + q \ .$$

for some constant q.

Assumption 4. *For a fixed $x \in \mathsf{X}$, suppose that:*

i. $B(x)$ *is a compact subset of* B.

ii. $c(x, \cdot)$ *is lower semi-continuous on* $B(x)$.

iii. The mapping $b \mapsto Ph$ *is continuous for all bounded measurable* h.

iv. The mapping $b \mapsto PV$ *is continuous.*

v. $s(x, \cdot)$ *is continuous on* $B(x)$.

The next theorem establishes the existence of stationary AC-optimal policies.

Theorem 9. *[89] If Assumptions 3 and 4 hold, then:*

1. *There exist a constant* $\varrho^* \geq 0$, *a continuous function* $\phi^* \in L_\infty^V$ *and a stationary policy* $\pi^* \in \Pi_0$ *such that the triplet* $(\varrho^*, \phi^*, \pi^*)$ *satisfies the average cost optimality equation (ACOE):*

$$\varrho^* + \phi^*(x) = \min_{b \in B(x)} \left[c(x, b) + \int_{\mathsf{X}} \phi^*(y) P(dy|x, b) \right] \tag{2.9}$$

$$= c_{\pi^*}(x) + P_{\pi^*}\phi^*(x), \quad \forall x \in \mathsf{X} \ ; \tag{2.10}$$

2. π^* *is AC-optimal and* ϱ^* *is the optimal AC-function.*

3. *The functions* ϕ_π *in Theorem 8 and* ϕ^* *satisfy*

$$\phi^* = \inf_{\pi \in \Pi^0} \phi_\pi \ .$$

Conversely, standard dynamic programming arguments show that any triplet (ϱ, ϕ, π) satisfying the ACOE with $\phi \in L_\infty^V$ has π as an AC-optimal policy and ϱ^* as the optimal AC-function provided that $1/N \, \mathrm{E}_x^{\pi^*}[\phi(\mathbf{x}(N))] \to 0$ as $N \to \infty$ [6, Thm. 5.1].

Analogously, we can define the average cost optimality inequality (ACOI):

$$\begin{aligned}\varrho^* + \phi^*(x) &\geq \min_{b\in B(x)} \left[c(x,b) + \int_{\mathsf{X}} \phi^*(y) P(dy|x,b)\right] \\ &= c_{\pi^*}(x) + P_{\pi^*}\phi^*(x), \quad \forall x \in \mathsf{X} \ . \end{aligned} \tag{2.11}$$

If the ACOI is satisfied, then we have an upper bound on the cost provided by the policy π^*:

$$J(\pi^*, x) \leq \varrho^* \ .$$

The proof of Theorem 9 suggests the following value iteration algorithm to solve the ACOE . Let ϕ_N and ϱ_N be defined as follows:

$$\begin{aligned}\bar{\phi}_N(x) &:= \min_{b\in B(x)} \left[c(x,b) + \int \phi_{N-1}(y) P(dy|x,b)\right] \\ &=: c_{\pi_N} + P_{\pi_N}\phi_{N-1} \\ \varrho_N &:= \nu(\bar{\phi}_N) \\ \phi_N &:= \phi_N - \varrho_N \ , \end{aligned}$$

where $\phi_0 = \bar{\phi}_0 \equiv 0$. The value iteration algorithm is said to converge if

$$\varrho_N \to \varrho^* \text{ and } \phi_N \xrightarrow{L_\infty^V} \phi^* \text{ as } N \to \infty \ .$$

Theorem 10. *Under Assumptions 3 and 4, the above value iteration algorithm converges. Moreover, the convergence rate is at least exponential.*

Proof. One can use the same arguments as in the proof of Theorem 3.5 of [89] to show that the mapping $\phi_N \mapsto \phi_{N+1}$ is a contraction on L_∞^V, which establishes exponential convergence of ϕ_N by the Banach fixed point theorem (recall that a fixed point exist by the previous theorem). Since ν is a bounded functional on L_∞^V, the cost function ϱ_N also converges with exponential rate. □

We can also use an alternative value iteration as follows. Let s_N (which can be seen as a N-th stage cost) and π_N be defined as follows:

$$s_N(x) := \min_{b\in B(x)} \left[c(x,b) + \int s_{N-1}(y) P(dy|x,b)\right] \ ,$$

where $s_0 \equiv 0$. Let $z \in \mathsf{X}$ be an arbitrary but fixed state. Define a sequence of constants j_N and a sequence of functions $\phi_N(x)$ as

$$j_N := s_N(z) - s_{N-1}(z) \quad \text{and} \quad \phi_N^0(x) := s_N(x) - s_N(z) \ .$$

Then, this value iteration algorithm is said to converge if

$$j_N \to \varrho^* \text{ and } \phi_N^0(x) \to \phi^*(x) \text{ as } N \to \infty$$

for some ϕ^* and ϱ^* satisfying the ACOE.

This value iteration algorithm also converges under the conditions of Theorem 10. To see why, note that the convergence of ϕ_N implies the convergence of ϕ_N^0 since these two functions only differ by a constant.

Other solution methods involve approximation by linear programs [43] and policy iteration. More general conditions for the existence of optimal policies and for the solutions to the ACOE are discussed in [35, Chap. 10].

2.4.1 Multi-objective MDPs

Often in MDPs one seeks to optimize conflicting objectives such as control performance and control effort. For a set of nonnegative cost functions $\{c_i\}_{i=1,N}$, this is typically done by minimizing the weighted cost $c_\lambda = \sum_i \lambda_i c_i$, where $\lambda = [\lambda_i] \geq 0$, and then the Pareto frontier can be constructed by varying λ. A problem that we face with this approach in this monograph is that the Pareto frontier may contain points that are not realized by deterministic policies. In this section we show how to calculate the missing points on the Pareto frontier from the deterministic policies. In general, constructing the Pareto frontier with a certain precision by searching on the weights λ may be computationally expensive. Here we assume that the frontier is smooth enough so that it can be well approximated with a small set of weights λ.

Denote by $J_i(\pi)$ the average c_i cost achieved by the policy π. For simplicity, we assume that J_i does not depend on the initial condition x. The *Pareto set* $\mathcal{P}$ is

the set of achievable cost combinations, i.e., $\mathcal{P} = \{z \in \mathbb{R}^N_{\geq 0} : z(i) = J_i(\pi), \pi \in \Pi\}$. If we adopt element-wise inequality as a partial order relation in $\mathcal{P}$, the *Pareto frontier* $\partial\mathcal{P}$ is defined as the set of minimal (infimal) elements of $\mathcal{P}$.

Proposition 3. *Suppose that Assumptions 3 and 4 hold for every cost c_i. Then, an AC-optimal stationary policy π^* for the cost c_λ exists for every weight $\lambda \geq 0$ and satisfies $[J_i(\pi^*)] \in \partial\mathcal{P}$.*

Proof. Let $J_\lambda(\pi)$ be the average c_λ cost achieved by policy π. Since lim sup is subadditive, we have that

$$J_\lambda(\pi) \leq \sum_{i=1}^{N} \lambda_i J_i(\pi), \quad \pi \in \Pi \ . \tag{2.12}$$

This implies that λ defines a supporting hyperplane for $\mathcal{P}$ passing through $[J_i(\pi^*)]$, where the existence of AC-optimal policy π^* for c_λ is given by Theorem 9. By Assumption 3, Theorem 8 and Theorem 4, we have that the limits in the definition of the average costs J_i exist and, therefore, that (2.12) holds with equality for π^*. Therefore, since $\lambda \geq 0$, $[J_i(\pi^*)] \in \partial\mathcal{P}$. □

We will show next that $\partial\mathcal{P}$ is a convex curve, which implies that we can find all points in $\partial\mathcal{P}$ by solving the c_λ-average cost problem and varying λ. One problem with this procedure is that we cannot find the optimal policies corresponding to the parts of $\partial\mathcal{P}$ that are not strictly convex. Indeed, our weighted cost procedure may return only one optimal policy, even when the supporting hyperplane defined by λ intersects $\partial\mathcal{P}$ at more than one point. By convexity of $\partial\mathcal{P}$, this intersection defines a convex set. It is not hard to see that the extreme points of this set can be realized by stationary policies. Indeed, we can see geometrically that there exist supporting hyperplanes that intersect $\partial\mathcal{P}$ only at these extremes. It turns out that the points on this convex intersection can be achieved by policies that randomize among the stationary policies that correspond to the extreme points of this set.

We say that $\partial\mathcal{P}$ is convex if it is the minimal set of its convex hull.

Proposition 4. *Suppose that Assumptions 3 and 4 hold for every cost c_i. Then, the Pareto frontier $\partial\mathcal{P}$ is convex. Moreover, every point in $\partial\mathcal{P}$ can be achieved by either a stationary policy or a randomized policy of the form*

$$\pi_{b|x} = \Pr\{\mathbf{b}(n) = b \mid \mathbf{x}(n) = x\} = \sum_{j=0}^{N} w_j(x)\delta_{\pi_j(x)}(b) \ , \tag{2.13}$$

where π_j are stationary policies such that $[J_i(\pi_j)] \in \partial\mathcal{P}$ and

$$w_j(x) = \begin{cases} \frac{\alpha_j d\mu_j}{d(\sum_{j=0}^{N}\alpha_j\mu_j)}, & \text{for } x \in \operatorname{supp}\sum_{j=0}^{N}\alpha_j\mu_j \\ 1/N, & \text{else} \end{cases} \ , \tag{2.14}$$

where $\alpha_j \geq 0$, $\sum_{j=0}^{N}\alpha_j = 1$ and μ_i are the invariant measures corresponding to π_j.

Proof. By Proposition 3, every supporting hyperplane to the set $\{z \in \mathbb{R}^N : \exists y \in \partial\mathcal{P} \text{ such that } z \geq y\}$ can be defined by a normal $\lambda \geq 0$ and a point $[J_i(\pi^*)]$, where π^* is a stationary AC-optimal policy for the cost c_λ. Let π_i be N of these stationary policies that define supporting hyperplanes. By linearity of P, $\mu_x := \sum_{i=1}^{N}\alpha_i\mu_i$ is the invariant measure when the policy $\pi_{b|x}$ is executed. The resulting process is Markov and ν-irreducible since its probability kernel randomizes between ν-irreducible kernels. Then, by Assumption 3, Theorem 8 and Theorem 4, we conclude that

$$J_j(\pi_{b|x}) = \int_{\mathsf{K}} c_j(x,b)\mu_x(dx)\pi_{b|x}(db) = \sum_{i=0}^{N}\alpha_i J_j(\pi_i)$$

for $j = 1, \ldots, N$, where the second equality follows from (2.13) and (2.14). This implies that the the convex hull of the of the average costs achieved by the policies π_i is in $\mathcal{P}$. Noting that the convex hull of a set is the same as the convex hull of the points in the set that admit a supporting hyperplane, we conclude that $\partial\mathcal{P}$ is a subset of its convex hull and, hence, convex. From this we have that every point on $\partial\mathcal{P}$ admits a supporting hyperplane and, therefore, this point is achieved by either a stationary policy or a randomized policy of the form $\pi_{b|x}$. □

Chapter 3

Jump Control of Probability Densities

In this chapter, we propose to control the probability density of a process using 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. In addition, we provide general results that have wide application in the study of ergodicity in PDPs, beyond the specific control design problem addressed here.

3.1 Problem Description

In this chapter we consider the PDP model of Section 2.3. In our setting, the vector field $f : \mathsf{X} \times \mathbb{M} \to \mathsf{X}$ is given and $\mathbf{m}(t)$ should be viewed as a control variable. The controller cannot measure the state $\mathbf{x}$ directly; instead, an observation variable $\mathbf{z}(t) = q(\mathbf{x}(t))$ is given. In general, the function $q(x)$ is not known to the controller, which has access to $\mathbf{z}(t)$ only.

Assuming that $q(x)$ is nonnegative and $\int_{\mathsf{X}} q(x)\ \ell(dx) < \infty$, our objective is to design λ and Q such that a randomized controller will select $\mathbf{m}(t)$ as a function of the observations $\{\mathbf{z}(\tau); 0 \leq \tau \leq t\}$ collected up to time t so that the marginal

probability density $\int_{\mathbb{M}} p(x,m,t)\ \nu(dm)$ converges to $cq(x)$, where c is a normalizing constant chosen so that $cq(x)$ integrates to one. We shall see later that the knowledge of the normalizing constant c is not necessary to implement the proposed control law.

Although jumps in the state x are allowed in a general PDP, we restrict Q so that jumps in only m are allowed. Further, we restrict $Q(\xi,\cdot)$ to be absolutely continuous with respect to $\delta_x \times \nu$. For this jump kernel, it is convenient to define the *jump pdf* T_x as

$$Q((x,m),\{x\}\times B) = \int_B T_x(m',m)\ \nu(dm'), \quad (x,m)\in \mathsf{X}\times\mathbb{M}, B\in\mathcal{B}(\mathbb{M})\ ,$$

where $T_x(\cdot,m)$ is a probability density function with respect to the reference measure ν.

In practice, $q(x)$ may be itself a function of some real measurements $F(x)$. For example, we can select $q(x) = Z(F(x))$, where the function $Z(\cdot)$ is a design parameter used to guarantee that $Z(F)$ is nonnegative and integrable. The function $Z(\cdot)$ 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 $Z(\cdot)$ that leads to a nonnegative integrable function is

$$Z(F) = \begin{cases} F & \text{, if } F > \delta \\ \delta e^{F-\delta} & \text{, if } F \le \delta, \end{cases} \tag{3.1}$$

for some $\delta > 0$. Alternatively, if one is mainly interested in achieving a high density close to the maxima of $F(x)$, a possible choice for $Z(\cdot)$ is given by

$$Z(F) = F^n\ ,$$

for some $n > 1$, provided that F^n is already nonnegative and integrable [if not, one could also use Z to achieve this, as it was done in (3.1) above]. The well-known optimization method of simulated annealing arises from a similar objective when n is increased to infinity along consecutive iterations [88].

3.2 Controller Existence and Design

In this section we provide a family of control laws that achieve the first part of our objective: to make a given probability density a stationary density for the PDP. The second part, which is the convergence to the desired density, is covered in the next section.

Our main design tool is the Fokker-Planck-Kolmogorov equation (2.8), which is used 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 (2.8)] corresponds to an invariant marginal distribution $\int_{\mathbb{M}} p(x, m, t)\ \nu(dm)$ that is proportional to $q(x)$. 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 $q(x)$ has been scaled so that it is itself a probability density: $\int q(x)\ \ell(dx) = 1$. However, none of our results require this particular scaling.

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 $p(x, m, t) = h(x, m)$, $\forall (x, m) \in \mathsf{X} \times \mathbb{M}$ and $t > 0$, for a probability density $h(x, m)$ that integrates to 1: $\int_{\mathsf{X}\times\mathbb{M}} h(x, m)\ \ell(dx)\nu(dm) = 1$. We denote by $L^1(\mathsf{X} \times \mathbb{M})$ the space of real functions integrable with respect to $\ell \times \nu$. 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 11. *Given a continuously differentiable probability density $h(x, m) > 0$, $\forall (x, m) \in \mathsf{X} \times \mathbb{M}$, with $\nabla_x \cdot fh \in L^1(\mathsf{X} \times \mathbb{M})$, a necessary and sufficient condition for h to be an admissible invariant density for the PDP $(\mathbf{x}(t), \mathbf{m}(t))$ is given by*

$$\int_{\mathbb{M}} \nabla_x \cdot fh(x, m)\ \nu(dm) = 0,\ \forall x \in \mathsf{X}\ . \tag{3.2}$$

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

$$\lambda(x, m) = \frac{\alpha(x) - \nabla_x \cdot fh(x, m)}{h(x, m)}, \tag{3.3}$$

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 (2.8):

$$\nabla_x \cdot fh = -\lambda h + \int_{\mathbb{M}} T_x(m, m')\lambda(x, m')h(x, m')\nu(dm') \ . \tag{3.4}$$

Recall that, since $T_x(\cdot, m')$ is a pdf, $\int_{\mathbb{M}} T_x(m, m')\nu(dm) = 1$. Using this fact, condition (3.2) is obtained by integrating both sides of (3.4) 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

$$\lambda h = \alpha(x) - \nabla_x \cdot fh \tag{3.5}$$

as in (3.3). For this jump rate to be consistent, we need to have $\lambda \geq 0$, which is possible if one takes $\alpha(x) \geq \max_{m\in\mathbb{M}} |\nabla_x \cdot fh|$. Note that this quantity is bounded since $\mathbb{M}$ is compact and fh is continuously differentiable. In addition, since $\nabla_x \cdot fh \in L^1(\mathsf{X} \times \mathbb{M})$ and $\mathbb{M}$ is compact, this choice of α guarantees that $\lambda h \in L^1(\mathsf{X} \times \mathbb{M})$. Finally, provided that λh is integrable, we can replace (3.5) and T_x in (3.4) to conclude from Theorem 7 that h is indeed an invariant density for our choice of the pair (λ,T_x). □

Remark 1. *It may happen that the λ given by (3.3) 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 $|\nabla_x \cdot fh| \leq Mh$, $\forall(x, v)$.*

Remark 2. *The control law in Theorem 11 admits a straightforward generalization: one can verify that the conclusions also hold when the jump pdf satisfies $T_x > 0$, $\int_{\mathbb{M}} T_x(m, m')\nu(dm') = 1$, $\int_{\mathbb{M}} T_x(m, m')f(x, m')\nu(dm') = 0$ and $\int_{\mathbb{M}} T_x(m, m')\nabla_x \cdot f(x, m')\nu(dm') = 0$.*

Condition (3.2) may be quite restrictive on the vector field f. Since q is not known in advance, we need (3.2) to hold independently of q. If, however, we

allow $h(x, m)$ to be arbitrary, the only vector field f that satisfies (3.2) 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 $h(x, m) = \beta(m)q(x)$, $\forall(x, m)$, 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 (3.2) to obtain the following corollary.

Corollary 2. *Consider continuously differentiable probability densities h that can be factored as $h(x, m) = \beta(m)q(x) > 0, \forall(x, m) \in \mathsf{X} \times \mathbb{M}$, where $\beta > 0$ and $q > 0$ satisfy $\beta\nabla_x \cdot fq \in L^1(\mathsf{X} \times \mathbb{M})$. Then, a necessary and sufficient condition for all h of this form to be admissible invariant densities is given by*

$$\int_{\mathbb{M}} f(x, m)\ \beta(m)\nu(dm) = 0,\ \ x \in \mathsf{X}\ . \tag{3.6}$$

□

Remark 3. *The existence condition (3.6) may be restrictive for some dynamical systems since it requires the ability to "reverse" the vector field. This is a problem for systems with relative degree larger than zero. For example, consider the case in which $\mathsf{X} = \mathbb{R}^2$, $\mathbb{M} = [-1, 1]$, ν is a uniform probability measure and $f(x_1, x_2, m) = [x_2\ m]^T$. This PDP does not satisfy the existence condition (3.2) with $h(x, m) = \beta(m)q(x)$, $\forall(x, m)$. In this case, one would be interested in achieving some m-dependent invariant density $h(x, m)$ such that $\int_{\mathbb{M}} h(x, m)\nu(dm) = q(x)$. However, it is not clear whether that can be done using output feedback.*

3.2.1 Output Feedback Controller

We consider now the amount of information that is needed to implement the control law proposed in Theorem 11. Corollary 2 is especially useful because the condition in (3.6) does not depend on the function q, which is not known in advance. We will therefore choose $h(x, m) = \beta(m)q(x)$. 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 (3.3), which we can rewrite as

$$\lambda = \eta - f \cdot \nabla_x \ln q - \nabla_x \cdot f \ , \tag{3.7}$$

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

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

$$f \cdot \nabla_x \ln q(\mathbf{x}(t)) = \frac{d \ln q}{dt^+}(\mathbf{x}(t)) \ , \tag{3.8}$$

  where '+' denotes the derivative from the right. Therefore, the controller only needs to have access to the time derivative of the observed output $\mathbf{z}(t) = q(\mathbf{x}(t))$ 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 $\eta(x) = \alpha(x)/q(x)$, 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 $\eta \geq |f \cdot \nabla_x \ln q + \nabla_x \cdot f| = |q^{-1} \nabla_x \cdot fq|$.

In particular, when there exists some function $\phi : \mathbb{M} \to \mathbb{R}$ such that $\nabla_x \cdot f(x,m) = \phi(q,m), \forall (x,m)$, and a function M that satisfies $M(q) \geq \max_{m \in \mathbb{M}} |q^{-1} \nabla_x \cdot fq|$, we can use the following output feedback realization of the jump rate:

$$\lambda(\mathbf{x},\mathbf{m}) = M(\mathbf{z}) - \frac{d \ln \mathbf{z}}{dt^+} - \phi(\mathbf{m}) \ . \tag{3.9}$$

3.2.2 Implementation Details

Assume, for simplicity, that M is constant and that $\nabla_x \cdot f = 0$. According to (2.6), the probability of the process maintaining the same mode in the interval

$[0,t]$ is given by

$$\exp\left(-\int_0^t \lambda(\mathbf{x}(\tau),\mathbf{m}(\tau))d\tau\right) = \exp\left(-\int_0^t M - \frac{d}{d\tau}(\ln q(\mathbf{x}(\tau)))d\tau\right) = e^{-Mt}\frac{q(\mathbf{x}(t))}{q(\mathbf{x}(0))} \quad . \tag{3.10}$$

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

$$\mathbf{z}(t) \leq \mathbf{r}\ e^{M(t-\boldsymbol{T}_k)}\ \mathbf{z}(\boldsymbol{T}_k), \quad t \geq \boldsymbol{T}_k \quad . \tag{3.11}$$

As opposed to what (3.9) seems to imply, one does not need to take derivatives to implement the jump rate. Also, the control law is not changed if a constant scaling factor is applied to $q(x)$, which is important because we cannot apply a normalizing constant to the unknown function q.

Often physical quantities propagate with spatial decay not faster than exponential, which allows for the uniform boundedness of $\|\nabla_x \ln q\|$ and the existence of a constant M. If, however, the measured quantity has a faster decay rate, it may still be possible to achieve boundedness of $\|\nabla_x \ln q\|$ by preprocessing the measurements (as explained in Section 3.1) as long as a finite bound for their decay rate is known.

In addition, the constant M may be identified on the run. This can be done, for example, with the following update rule:

$$\mathbf{M}(t) = \begin{cases} |d(\ln \mathbf{z})/dt^+| + \epsilon, & \text{if } \mathbf{M}^-(t) \leq |d(\ln \mathbf{z})/dt^+| + \epsilon/2 \\ \mathbf{M}^-(t), & \text{else} \end{cases} \tag{3.12}$$

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

3.3 Ergodicity of the Controlled Process

In this section we investigate whether the above control strategy makes the probability density of the PDP converge to q as time goes to infinity. We summarize the results in this section with Theorem 12, which gives necessary and sufficient conditions for convergence.

Let $B_r(x)$ denote the open ball with radius r centered at $x \in \mathsf{X}$. We say that the system $\dot{x} = f(x,u)$, $u \in \mathbb{M}$, is *approximately controllable* if, for every $x_0, x_1 \in \mathsf{X}$ 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 12. *Suppose that $q > 0$ is a continuously differentiable density for which there exists a continuous and bounded function M and $\epsilon > 0$ satisfying $|\nabla_x \cdot fq|/q + \epsilon \le M(q)$. Consider the PDP $(\mathbf{x}(t), \mathbf{m}(t))$ with the output feedback control:*

$$T_x(\cdot,\cdot) \equiv 1, \quad \lambda(\mathbf{x},\mathbf{m}) = M(\mathbf{z}) - \frac{d \ln \mathbf{z}}{dt^+} - \phi(\mathbf{m}) \ . \tag{3.13}$$

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

i. $\int f(x,m)\ \nu(dm) = 0$, $\forall x \in \mathsf{X}$;

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

Moreover, when the above convergence holds, the PDP is an aperiodic Harris recurrent process and the following convergence of empirical averages holds: for every $\tau > 0$ and every ψ such that $\psi g \in L^1(\mathsf{X} \times \mathbb{M})$,

$$n^{-1} \sum_{k=0}^{n-1} \psi(\mathbf{x}(\tau k), \mathbf{m}(\tau k)) \to \int_{\mathsf{X}\times\mathbb{M}} \psi(x,m) q(x)\ \ell(dx)\nu(dm) \quad a.s. \tag{3.14}$$

for all initial conditions. □

The proof of this theorem will appear later in the chapter. Before that we discuss the assumptions and conclusions of the theorem.

Remark 4. *We say that a set* $\mathsf{F} \subset \mathsf{X}$ *is a* positive basis *if* 0 *is in the (algebraic or topological) interior of the convex hull of* F*. A typical case in which condition* (ii) *is satisfied is when* $\{f(x,m); m \in \mathbb{M}\}$ *contains a positive basis for* X *for all* $x \in \mathsf{X}$ *(see the Filippov-Wazewski argument in Proposition 6). We can always define a reference measure* ν *to satisfy condition* (i) *as long as the positive basis behaves uniformly with* x*.* □

Next, we show that convergence is preserved if M is identified on the run using (3.12).

Theorem 13. *Suppose that* $\sup_{\mathsf{X}\times\mathbb{M}} |\nabla_x \cdot fq|/q < \infty$ *and that* $\nabla_x \cdot f = 0$*. Then, the conclusions in Theorem 12 hold when the jump rate is replaced by*

$$\lambda(\mathbf{x}, \mathbf{m}) = \mathbf{M}(t) - \frac{d \ln \mathbf{z}}{dt^+} \ , \tag{3.15}$$

where $\mathbf{M}(t)$ *is identified on the run using (3.12).*

Proof. Let $\bar{M} = \sup_{\mathsf{X}\times\mathbb{M}} |\nabla_x \cdot fq|/q$. From (3.12), $\mathbf{M}(t)$ increases by at least $\epsilon/2$ at every update. Thus, $\mathbf{M}(t)$ must achieve a limit $\mathbf{M}_0 \leq \bar{M} + \epsilon$ in finite time. Suppose $\mathbf{M}_0 \leq \bar{M}$ and let $C_0 = \{(x,m) \in \mathsf{X} \times \mathbb{M} : |f \cdot \nabla_x \ln q| + \epsilon/2 < \mathbf{M}_0\}$. This definition implies that $(\mathbf{x}(t), \mathbf{m}(t)) \in C_0$ for all time. Let q_0 be a probability density such that $q_0 = q$ on C_0 and $|f \cdot \nabla_x \ln q_0| + \epsilon/2 < \mathbf{M}_0$. Since $\lambda \geq 0$ on C_0, we can apply Theorem 12 to conclude that $p(x,v,t)$ converges to q_0. But, since $\mathbf{M}_0 \leq \bar{M}$,

$$\int_{C_0^c} q(x)\ell(dx)\nu(dm) \geq \int_{\{|f\cdot\nabla_x \ln q|+\epsilon/2>\bar{M}\}} q(x)\ell(dx)\nu(dm) > 0 \ , \tag{3.16}$$

where the last inequality follows from the continuity of $f \cdot \nabla_x \ln g$. This contradicts the convergence of $p(x,v,t)$ to $q_0(x)$ since $\int_{C_0} q_0(x)\ell(dx)\nu(dm) < 1$. Therefore, $\mathbf{M}(t)$ achieves a limit $\mathbf{M}_0 > \bar{M}$ in finite time and Theorem 12 can be applied to conclude convergence after the limit is reached. □

3.3.1 Ergodicity for the PDP

In this section we derive some new results regarding the ergodicity of invariant measures of PDPs. While some ergodicity results specific for PDPs may be found

in the literature (see e.g. [23]), those rely mostly on Foster-Lyapunov criteria and do not appear to be suited our purposes, 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 $(x,m) \in \mathsf{X} \times \mathbb{M}$ and any set B with $\nu(B) > 0$, there is a positive probability that $\{x\} \times B$ will be eventually reached from (x,m).

Assumption 5. *i. the jump Markov chain is mode-irreducible.*

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

$$(x,m) \mapsto \int_{\mathbb{M}} T_x(m',m)\psi(x,m')\ \nu(dm') \tag{3.17}$$

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.*

□

Proposition 5. *Suppose that Assumption 5 (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.,*

$$P^{t_1}\big((\bar{x}(0), \bar{m}(0)), B_{\epsilon_1}(\bar{x}(t_1), \bar{m}(t_1))\big) > 0,\ \forall t_1, \epsilon_1 > 0. \tag{3.18}$$

□

Proof. By Assumption 5 *(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 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 5 *(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 6. *Suppose that Assumption 5 (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.,*

$$P^t\big((x_0, m_0), B_{\epsilon_1}(\hat{x}(t), m_1)\big) > 0, \ \forall t, \epsilon_1 > 0 \tag{3.19}$$

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 2), we can apply the Filippov-Wazewski theorem [8, 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 2 *(i)*, we can apply Theorems 2.20 and 2.24 of [40] 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 5, 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 5, 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. □

Proposition 7. *Under Assumption 5 (i)-(iv), the PDP* $(\mathbf{x}(t), \mathbf{m}(t))$ *is an aperiodic* $\ell \times \nu$*-irreducible T-process.* □

Proof. By Proposition 6, $\ell \times \nu$-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 6 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 [25, Thm. 27.6], Assumption 5 (*ii*) implies that the PDP has the (weak) Feller property as defined in Chapter 2. Given the Feller property, the $\ell \times \nu$-irreducibility and the fact that $\text{supp}(\ell \times \nu) = \mathsf{X} \times \mathbb{M}$ has non-empty interior, we can use Proposition 1 to conclude that the PDP is a T-process. □

Proposition 8. *Suppose that Assumption 5 (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.* □

Proof. Let μ denote the invariant measure corresponding to h. By Proposition 7, the PDP is an aperiodic $\ell \times \nu$-irreducible T-process. Therefore, the space $\mathsf{X} \times \mathbb{M}$ admits a decomposition into a maximal Harris set H with invariant measure μ and a transient set E as in Theorem 3. 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. □

Proof of Theorem 12. (Necessity) The necessity of condition *(i)* follows from Corollary 2. To see the necessity of condition *(ii)*, note that the convergence of $p(x,m,t)$ to $q(x)$ implies that the process is μ-irreducible, where $\mu(dx,dm) = q(x)\ell(dx)\nu(dm)$, which implies the controllability condition since $q > 0$.

(Sufficiency) It follows from Theorem 11 and Corollary 2 that q is an invariant density for the pair (λ, T_x) presented. To prove convergence, we apply Proposition 8. The inequality $|\nabla_x \cdot fq|/q + \epsilon < M$ implies that $\lambda = M - \nabla_x \cdot fq/q \geq \epsilon$ and

that λ is uniformly bounded. Therefore, Assumption 5 *(i)* holds since $\lambda \geq \epsilon$ and a uniform jump distribution imply mode-irreducibility. Since fq is continuously differentiable in x, we have that λ is continuous and, therefore, Assumption 5 *(ii)* holds. Assumption 5 *(iii)-(iv)* follows from conditions *(i)* and *(ii)*. Therefore, we have that the process is aperiodic positive Harris recurrent, which implies W-ergodicity and convergence in total variation by Theorem 4. Clearly, the same convergence result as in (2.4) 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 Theorem 2. □

3.4 Jump Control versus Diffusion Control

In the same way we can control probability densities with random jumps in the vector field, we can control probability densities using controls which undergo Brownian motion. Consider the process

$$d\mathbf{x}(t) = f(\mathbf{x}, \boldsymbol{\theta})dt \tag{3.20}$$

$$d\boldsymbol{\theta}_i(t) = \sigma_i(\mathbf{x}, \boldsymbol{\theta})d\mathbf{w}_i, \ i = 1, \ldots, n_\theta \tag{3.21}$$

where $[\boldsymbol{\theta}_i]$ is a vector, $\mathbf{w}_i(t)$ is the standard 1-dimensional Wiener process and σ_i are nonnegative functions of $\mathbf{x}$ and $\boldsymbol{\theta}$.

As before, we want to design σ in order to induce a stationary distribution for $\mathbf{x}(t)$. We can write the standard Fokker-Planck equation for this process as

$$\frac{\partial p}{\partial t} + \nabla_x \cdot fp = \frac{1}{2}\sum_{i=1}^{n_\theta} \frac{\partial^2 \sigma_i^2 p}{\partial \theta_i^2}$$

For simplicity, let $n_\theta = 1$ and $\nabla \cdot f = 0$. If we replace $p(x, \theta, t) = q(x)$ and solve for σ_i^2, we obtain

$$\sigma^2(x, \theta) = \eta + 2 \int\int f(x, \theta)\ d\theta d\theta \cdot \nabla \ln q$$

where η is some function independent of θ and the integrals are indefinite. With our output measurement model, only $f \cdot \nabla \ln q$ can be observed. This means that

the double integral must be an affine function of f in order that we can implement σ^2 with output feedback. This is only possible if the components f_j of the vector field are of the form

$$f_j = g_j(x)\cos(\omega\theta + \phi_j),\ j = 1, \ldots, d\ , \tag{3.22}$$

where ω and ϕ_j are constants and g_j is some function of the state.

This imposes an important restriction on the structure of diffusion-based controllers as compared to jump-based controllers. Conceivably, more general vector fields f could be considered if one added a drift term to (3.21) or if one considered $n_\theta > 1$. However, it is still an open problem whether the resulting control can be implemented with output feedback.

For vector fields of the form (3.24), we obtain an expression for the variance analogous to the expression for λ in the previous sections:

$$\sigma^2 = \eta - 2\omega^{-2} f \cdot \nabla_x \ln q\ . \tag{3.23}$$

As before, η must be such that σ^2 is nonnegative.

This result can be readily extended to the case $n_\theta > 1$ if f is of the form

$$f_j = g_j(x)\prod_{i=1}^{n_\theta}\cos(\omega_i\theta_i + \phi_{ji}),\ j = 1, \ldots, d\ , \tag{3.24}$$

where ω_i and ϕ_{ij} are constants and g_j is a function of the state x. The corresponding variance is

$$\sigma_i^2 = \eta_i - 2\kappa_i f \cdot \nabla_x \ln q\ , \tag{3.25}$$

where κ_i are any nonnegative weights satisfying $\sum_{i=1}^{n_\theta}\kappa_i\omega_i^2 = 1$.

To prove convergence, define the Kullback-Leibler divergence between two solutions p_1 and p_2 of the Fokker-Planck equation by

$$H(t) = \int_{\mathbb{R}^d\times\mathbb{R}^{n_\theta}} \ln\frac{p_1(x,\theta,t)}{p_1(x,\theta,t)/2 + p_2(x,\theta,t)/2}\ p_1(x,\theta,t)dxd\theta\ . \tag{3.26}$$

We can use the Fokker-Planck equation as in [76, Eq. 6.18] to compute the time derivative of H:

$$\dot{H}(t) = -\int_{\mathbb{R}^d \times \mathbb{R}^{n_\theta}} \left(\frac{\partial \ln R}{\partial \theta} \right)^2 p_1 \sigma^2 dx d\theta \ , \tag{3.27}$$

where $R = p_1(x, \theta, t)/(p_1(x, \theta, t)/2 + p_2(x, \theta, t)/2)$. Therefore, if p_1 and p_2 are two invariant probability densities, R must be independent of θ $\sigma^2 p_1$-a.e. Suppose that $\sigma^2 > 0$ for all (x, θ) and that also the stationary density is such that $q(x) > 0$. Then, if p is another invariant density, it must be independent of θ. If we divide the respective Fokker-Planck equations by p and q then subtract the equations, we obtain that the two invariant densities must satisfy

$$f \cdot \nabla_x \ln \frac{p}{q} = 0, \ \forall (x, \theta) \ . \tag{3.28}$$

If $\text{span}\{f(x, \theta); \theta \in \mathbb{R}^{n_\theta}\} = \mathbb{R}^d$ for all x, (3.28) implies that $p \equiv q$. With this we conclude that the invariant density is unique and therefore ergodic. This allows us to apply the continuous analog of Theorem 1 to conclude convergence of empirical averages.

In summary, the empirical density of $\mathbf{x}(t)$ converges weakly to $q(x)$ provided that $\text{span}\{f(x, \theta); \theta \in \mathbb{R}^{n_\theta}\} = \mathbb{R}^d$ and $\sigma^2 > 0$ for all x.

As we have seen, this diffusion control strategy shares many similarities with our jump control. Its main drawback is that it requires a special topology on the space of controls θ so that they can diffuse. A consequence of that is the restrictive form required for f in (3.24).

3.5 Comments and Open Problems

We have presented a Markov Chain Monte Carlo approach for dynamical systems. As discussed in Remark 3, the main limitation of our solution is given by the necessary condition (3.6), which prevents it to work with systems whose relative degree is higher than or equal to one. At the same time, we have seen that (3.6) implies the aperiodicity of the underlying process. This suggests that possible

solutions for the case of systems with higher relative degrees may have to be periodic.

One major challenge in further understanding the jump control problem is in the difficulty of treating partial integro-differential equations analytically. Analytical solutions can no longer be found if we consider slight modifications of the proposed model such as introducing disturbances. A possible way to deal with this problem is to use perturbation methods for semigroups or their resolvents. To this purpose, knowing Lyapunov functions for the unperturbed process is fundamental.

In the spirit of Markov Chain Monte Carlo, it would be valuable to explore the use of simulated annealing, for application in optimization problems, and the use of optimal stopping rules that decide when enough information about q has been collected.

It is also possible to envision the use of this technique in a parallel computing implementation of Markov Chain Monte Carlo. Suppose that $q(x)$ is some function hard to compute, but whose computation can be parallelized. For example, if $q(x) = \mathrm{E}_\theta\, q_\theta(x)$ and computing q_θ is computationally expensive, the computation can be parallelized by having different nodes of a computer cluster to compute samples q_{θ_i} and then communicate the result to a central node that computes $q(x) \approx \sum_i q_{\theta_i}(x)$. The benefit of using jump control is in that, assuming that the nodes random number generators are synchronized, the vector x does not need to be transmitted to the nodes but only the decision of whether to jump or not, which represents time savings in the case x is very large. However, the same effect could be achieved using the Metropolis-Hastings algorithm, which can also be implemented without explicitly transmitting x. Hence, it is not clear if there is a particular advantage in using jump control.

Chapter 4

Applications of Jump Control to Mobile Robotics

This chapter addresses the application of jump control of probability densities in the area of mobile robotics. In this type of application, the state $\mathbf{x}$ typically includes the position of a mobile robot that can take point measurements $\mathbf{z} = q(\mathbf{x})$ at its current location. We discuss three main applications: monitoring, deployment and search.

The control algorithms proposed here are motivated by the chemotactical motion of the bacterium *E. coli*. Due to its reduced size, the *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* [14, 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 [14] 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 maximum of the sensed quantity.

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 the *E. coli*. This bacterium also does not have access to global or local 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 inspired us to formulate the problem of jump control of probability densities and apply it to the field of mobile robotics.

While mimicking chemotaxis is not a new solution to optimization problems, see e.g. [45, 91, 71, 28, 59], our control strategy is distinct in that we are able to provide formal statements about the stationary density and the convergence to it.

4.1 Applications in Mobile Robotics

4.1.1 Monitoring Applications

In *monitoring applications*, one attempts to estimate a spatially-defined function q using mobile robots. Potential applications for this work thus include chemical plant safety, hydrothermal vent prospecting, pollution and environmental monitoring, fire or radiation monitoring, etc.

Traditional solutions to this problem involve sweeping a grid on the space. In our solution, we use jump control of probability densities to make the spatial

distribution of a group of robots coincide with q. Then, we have an external observer to estimate q by keeping track of the positions of robots. It is also possible to have agents transmit their measurements to said observer.

It is well known in the MCMC literature that a random walk that samples the space according to a probability density q is more efficient to describe q than a uniform grid since points are sampled with a probability that is proportional to their importance to the function being described. Another advantage of our solution with respect to grid methods is that it does not require a complicated planning, which would involve agent coordination, high memory requirements and position sensors.

The convergence of empirical averages provides the basis for a procedure to estimate $q(x)$ by observing the positions $\mathbf{x}_n$ of N identical vehicles performing the above jump control strategy. We start by partitioning the region of interest into a family of sets $\{A_i \subset \mathsf{X}\}$, 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 $q(x)$ on the set A_i. To see why this is the case, we define

$$G_{N,M}(A_i) = \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{k=0}^{M-1} 1_{A_i}(\mathbf{x}_n(k\tau)) \ . \tag{4.1}$$

Assuming that the vehicles have mutually independent motion, we can use ergodicity as in Theorem 12 to conclude that

$$G_{N,M}(A_i) \to G(A_i) := \int_{A_i} q(x) \ \ell(dx) \ \text{ a.s.} \tag{4.2}$$

as $M \to \infty$. This shows that $q(x)$ can be estimated by averaging the observations of the vehicles' position as in (4.1). The use of multiple independent agents ($N > 1$) improves the estimates according to the relation

$$\text{var}(G_{N,M}) = \frac{\text{var}(G_{1,M})}{N} \ . \tag{4.3}$$

4.1.2 Deployment Applications

In *deployment applications*, a group of robots is required to distribute themselves in an environment based on the values of the measured function q. These measurements may be, for example, 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. Alternatively, q may be some radio signal used to control the spatial density of agents.

Most deployment strategies in the literature require agents to converge to some static equilibrium [18]. This type of deployment offers the benefit of minimizing some coverage metric. However, it demands from agents the ability to communicate and measure their positions. Another possible disadvantage of a static deployment is that it may prevent agents from exploring other regions of the space.

We describe next in which sense our deployment strategy is optimal. We can translate the objective of making $p(x, m, t)$ converge in total variation to q as the maximization of a relevant reward function. From [58], this convergence implies that

$$H(t) = \int_{\mathsf{X}} \int_{\mathbb{M}} p(x, m, t) \ln \left(\frac{1}{2} + \frac{1}{2} \frac{q(x)}{p(x, m, t)} \right) \ \ell(dx)\nu(dm) \to 0 \ ,$$

where $H(t)$ is the Kullback-Leibler divergence between $p(x, m, t)$ and the convex combination $1/2\ q(x) + 1/2\ p(x, m, t)$. Since $H(t) \leq 0$ with equality to zero if and only if $q(x) = p(x, m, t)$ a.e., one can regard $H(t)$ as a cost functional that is being maximized by our control. More specifically, what is being maximized is the expected value of an increasing concave function of q/p:

$$H = \mathrm{E}_p \left[\ln \left(\frac{1}{2} + \frac{1}{2} \frac{q}{p} \right) \right] \ . \tag{4.4}$$

In a multi-agent scenario in which the number of agents is large enough so that the probability density $p(x, m, t)$ approximates the density of agents at point (x, m, t), we can interpret this as a maximization of the average intensity of the field q per vehicle.

Moreover, it is a well-known fact that H increases monotonically with time [57, Thm. 9.2.2], which means that agents steadily improve their deployment metric H.

4.1.3 Search Applications

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 $q(x)$ 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.

This objective is achieved by a method closely related to simulated annealing [88]. If $F(x)$ denotes the spatial field whose maximum we search for, we set q to

$$q(x) = Z(F(x)), \quad Z(F) = F^n \ ,$$

for some $n > 1$. For high values of n, we have that the vehicles spend most of the time on a neighborhood of the maximum of F. As we will see later, a drawback of using n too large is that convergence to the stationary density may be slow. In practice, a large n makes the search too sensitive to measurement noise.

On the other hand, a benefit of not using very large n is that vehicles may also find local maxima of the function F, which is beneficial in the source-seeking context since these maxima are often caused by secondary sources of the measured physical quantity.

In simulated annealing, the exponent n is gradually increased as time progresses so that $\mathbf{x}$ actually converges to the maximum of F.

We adopt the suggestive name of *optimotaxis* to designate this search procedure. An important feature of optimotaxis is that it can be used with a broad class

of signal profiles, including the ones with multiple maxima, a feature that is shared with a few other stochastic optimization algorithms which are not constrained by vehicle kinematics [3, 26].

4.2 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 [84], 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.

4.2.1 Optimotaxis

We have introduced the term *Optimotaxis* in [66] to designate our jump control strategy for a dynamics that resemble that of the tumble and run motion of the *E. coli*. We consider vehicles moving with position $x \in \mathsf{X} = \mathbb{R}^d$ and velocity v, where v belongs to the unit sphere $\mathbb{M} = \mathbb{S}^d$. The measure ν is the normalized surface measure on the sphere. In this case, the mode is represented by v and we have $f(x, v) = v$. Our objective is to make the probability density of the vehicles position converge to the observed function $q(x)$ and then have an external observer that can measure the vehicles position to collect information about $q(x)$.

It is straightforward to show that the existence condition (3.6) holds and that $\{f(x, v); v \in \mathbb{M}\}$ is a positive basis for all x (see Remark 4). Therefore, we can apply Theorem 12 and to obtain

$$T_x \equiv 1 \text{ and } \lambda(x, v) = \eta - v \cdot \nabla_x \ln q, \quad , \tag{4.5}$$

where $\eta > \|\nabla_x \ln q\|$.

Next, we present numerical experiments to illustrate the proposed optimization procedure. The desired stationary density is taken to be $q(x) = cF^n(x)$, 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. 4.1. 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. Yet, a portion of the agents clearly indicates the existence of the local maximum. This means that the information on secondary sources (local maxima) is not lost.

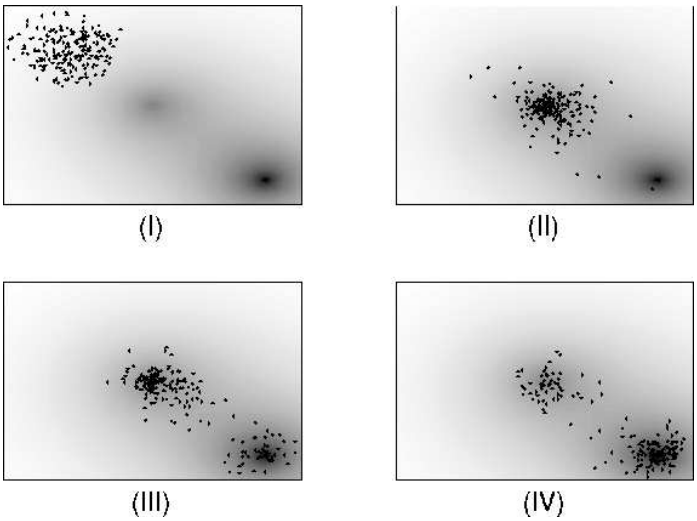

Figure 4.1: Different stages of optimotaxis in the presence of two maxima. Black dots represent agents position whereas the background intensity represents the signal intensity. $F(x) = 0.4e^{-\|x\|} + 0.6e^{-\|x-[1.5\ -1.5]'\|}$, $q(x) = F^n(x)$ with $n = 10$.

To quantify the convergence of the positions of the agents to desired distribution $q(x)$, we compute the correlation coefficient between the vectors $[G(A_i)]_i$ and

$[G_{M,N}(A_i)]_i$. This coefficient was calculated using a space grid with resolution 0.068 and its time evolution appears in Fig. 4.2.

Also included in Fig. 4.2 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 $q(x)$ to calculate the coefficient. Interestingly, the addition of noise does not seem to considerably affect 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 negative impact on convergence time.

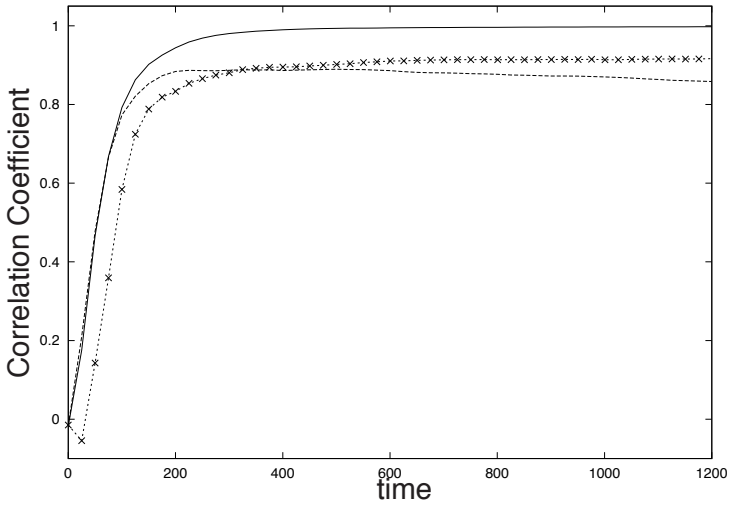

Figure 4.2: Time evolution of the coefficient of correlation.
The noiseless case (solid), the quantized measurements case (cross), and the exogenous noise case (dashed). The number of quantization levels is 64. The noise added to $\dot{v}$ is white Gaussian with standard deviation 10^{-2} along each 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. 4.1 for more details.

The sensitivity of the procedure with respect to the parameter n is studied in Fig. 4.3. 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 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.

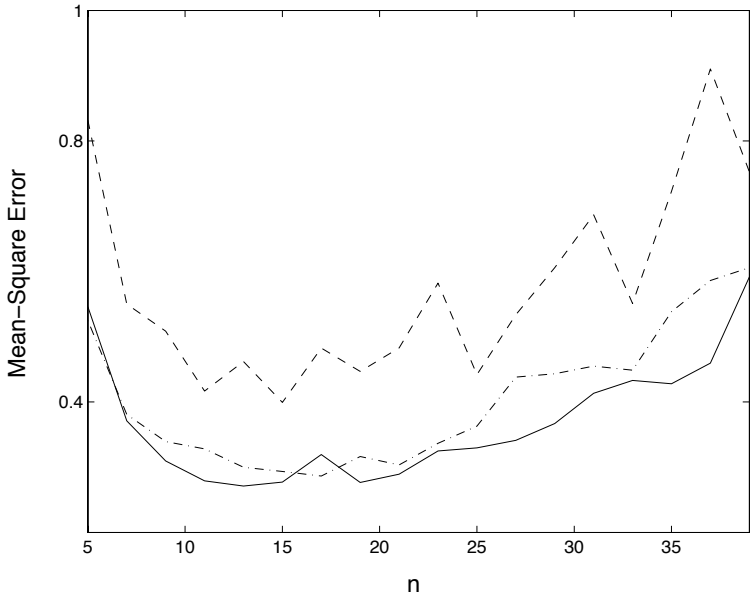

Figure 4.3: Mean-square error with respect to the maximum of $F(x) = e^{-\|x\|}$ as a function of n.
Noiseless case (solid), quantized $F(x)$ (dashed), and exogenous noise (dash-dotted). The number of quantization levels is 128. The noise added to $\dot{v}$ is white Gaussian with standard deviation 10^{-3} in each axis.

4.2.2 Reeds-Shep Car

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

$$f(x,v) = \begin{bmatrix} v_1 \cos\theta \\ v_1 \sin\theta \\ \omega \end{bmatrix} , \tag{4.6}$$

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 [84].

The vector field (4.6) satisfies the existence condition (3.6). Hence, we can use λ and T_x as in Theorems 11 and 12 to make q an invariant density. More precisely, $\lambda = \eta - f \cdot \nabla_x \ln q$ and $T_x \equiv 1$.

Even though $\{f(x,v);\ v \in \mathbb{M}\}$ does not contain a positive basis for X, it is still easy to verify the controllability condition in Theorem 12. Given x and y in X, 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). The controllability condition 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 X when $v_1 = v_0$ and $\omega = \omega_0$. 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_1, x_2) to (y_1, y_2) 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. Therefore, Theorem 12 gives the convergence of $p(x, m, t)$ to $q(x)$.

Figure 4.4 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 of the previous section 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 mentioning that in the case for which 0 linear velocity is not allowed convergence is only slightly slower than in the present case.

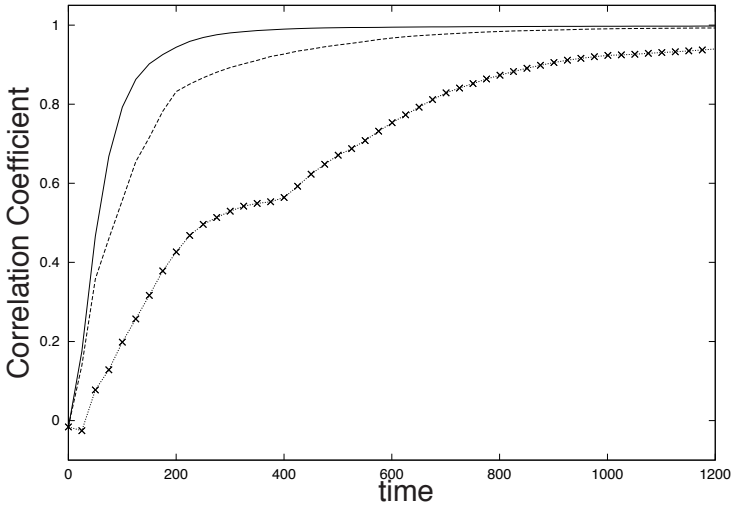

Figure 4.4: Evolution of the coefficient of correlation under turning constraints. Unconstrained turning case (solid), for the constrained turning case with $v_0 = 1$ and $\omega_0 = 0.3$ (dots) and for the constrained turning case with $v_0 = 1$ and $\omega_0 = 0.15$ (cross). The simulation setting is the same as of Figs. 4.1 and 4.2.

4.2.3 Navigation with beacons

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 12, we choose

$$\lambda = M - f \cdot \nabla \ln q - 2m_2 \ , \tag{4.7}$$

for some M sufficiently large. Note that f satisfies the hypotheses of Theorem 12: 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 $x\|\nabla \ln q\|$ 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

$$T_x(m, m') = \frac{1}{4} 1_{\mathbb{M}-\{m'_1 \times \{-1,1\}\}}(m) \ . \tag{4.8}$$

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 8 to conclude convergence to the invariant density q.

4.2.4 Chemotaxis as Search and Deployment Applications

Chemotaxis in the bacterium *E. coli* can be seen as a good example of how the jump control of probability densities is used in either search or deployment applications. In this section, we try to shed new lights on chemotaxis using our jump control framework.

At present, there is little understanding of the interplay between the intracellular dynamical system that describes signal transduction in *E. coli*'s chemotaxis and quantities such as the turning rate and turning distribution in a macroscopic, population-level description of motion. With two qualitative assumptions on the turning rate and and on the signal transduction mechanism, we can explain experimental observations and make conclusions about the stationary behavior of the population of bacteria.

Assumptions on The Macroscopic Model

A derivation of the macroscopic behavior of bacteria from their internal (intracelular) dynamics is proposed in [34]. Both tactic and kinetic responses involve two major steps: detection of the signal and its transduction into an internal signal that triggers the response. The following expression for the jump (tumbling) rate λ of a single bacterium was obtained in [34]) for the scalar case (Equation (6.36)):

$$\lambda = \eta - \frac{w(z(t))}{dt} \ , \tag{4.9}$$

where $\lambda_0 > 0$ is a constant, w is some function that models the internal dynamics in the cell and z is the concentration of chemoattractant sensed by the bacterium at time t. A similar expression was also derived in [2]. It was proposed in [95] that integral feedback is the mechanism that explains how time derivatives as in (4.9) are implemented. We further assume that the new orientation after tumbles is uniformly distributed on the sphere and that a bacterium velocity is constant in the run phase.

In the vast majority of the literature, it is the diffusion approximation of the above model, initially proposed in [54], that is used to model chemotaxis. One major limitation of this approximation is that it holds only for one-dimensional spaces. Because we do not use this approximation, our conclusions are valid for spaces with any dimension.

The Stationary Density

As before, let x be a space coordinate and let $F(x)$ denote the concentration of chemoattractant at position x, where we implicitly assume that the distribution of chemoattractant remains (approximately) constant in time. This would correspond to the consumption of nutrients occurring in a timescale that is much slower than chemotaxis. We denote by $p_t(x)$ the probability density of a single bacterium in space at time t. If the population of bacteria is large enough, we can take $p_t(x)$ to approximate the concentration of bacteria at point x and time t. With these assumptions, by Theorem 12 , the tumbling rate in (4.9) leads to a stationary density

$$p_\infty(x) = c_0 e^{w(F(x))} \ , \tag{4.10}$$

where c_0 is a normalizing constant.

Moreover, for a given stationary density $c_0 e^{w(F(x))}$, one can conclude from the Fokker-Planck-Kolmogorov equation that the only tumbling rate that achieves such a stationary density is the one given by (4.9) (up to an additive constant). *This suggests that a major reason why bacteria tumble according to (4.9) is that*

they seek some advantage from the ability of choosing the stationary density of the population. We will return to this point later in the section.

Hypothesis of Maximization of Sensitivity Range

In biophysics, the fact that perception in sensory systems has a logarithmic dependence on stimulus is well known as Weber's law. The fact that chemotaxis obeys Weber's law has been first investigated in [65] with inconclusive results. Weber's law was later confirmed experimentally in [16] by applying an attractant concentration with an exponential increase with time. More recently, [87] proposed a model to explain Weber's law in bacterial chemotaxis. *Here we explain Weber's law based on the simple hypothesis that maximal sensitivity range is a desired trait in bacteria.*

In oder to find regions with higher nutrient concentrations (or to remain in those regions), bacteria must be able to detect well a relative increment in chemoattractant levels. In addition, this ability is desired for both high or low chemoattractant levels. In other words, chemotaxis performance depends on the sensitivity of bacteria to a wide range of chemoattractant concentrations. We define the sensitivity $\mathcal{S}$ to chemoattractants as the change in the response w caused by relative changes in the stimulus z:

$$\mathcal{S}(z) = \left| z \frac{dw}{dz}(z) \right| \quad .$$

We can rewrite the expression for sensitivity as

$$\mathcal{S}(z) = \left| \frac{dw}{d \ln z}(z) \right| \quad .$$

We assume that a minimum sensitivity of α is necessary for good performance, where α may be determined by the level of disturbances such as rotational diffusion (on the other hand, measurement noise should put an upper bound on the sensitivity). Assume that z_0 is the maximum concentration that can be sensed by bacteria. Let us imagine, at first, that bacteria are free to choose w with the

only condition that it lies in the interval $[\underline{w}, \overline{w}]$. With such conditions, it is a simple exercise to verify that the functions w that maximize the range of effective sensitivity α are

$$w = \alpha \ln \frac{z}{z_0} + w_0 \ , \tag{4.11}$$

where w_0 is any constant satisfying $w_0 \geq \overline{w}$. *This result characterizes Weber's law as a direct consequence of the need to maximize sensitivity range.*

Next, we argue that (4.11) can be well approximated by bacteria even if they are not able to implement arbitrary functions w. To see this, let $\bar{z}$ denote the point of maximum of $\mathcal{S}(z)$. Then, $\bar{z}$ satisfies the first order condition

$$\frac{d\mathcal{S}}{d(\ln z)}(\bar{z}) = \frac{d^2 w}{d(\ln z)^2}(\bar{z}) = 0 \ .$$

Therefore, w as a function of $\ln z$ has zero curvature at the point of maximum sensitivity $\bar{z}$. This means that w is well approximated by an affine function of $\ln z$:

$$w(\ln z) = w(\bar{z}) + \bar{z} w'(\bar{z})(\ln z - \ln \bar{z}) + o((\ln z - \ln \bar{z})^3) \ . \tag{4.12}$$

A point of maximum sensitivity typically exists for w concave, in which case, we have from the first order maximality condition that $\bar{z} w'(\bar{z}) = -\bar{z}^2 w''(\bar{z}) > 0$. Therefore, any concave function w with a point of maximum sensitivity can be well approximated at the point of maximum sensitivity by $\alpha \ln z + b$, where $\alpha > 0$ and b are constants. In conclusion, (4.11) can be approximated as long as bacteria tune w so that the typical values of z are close to the point of maximum sensitivity.

In the references above, w is typically modeled as the fraction of cellular receptors occupied by molecules of the chemical attractant, which leads to the usual expressions found in cooperative binding:

$$w(z) = \frac{\kappa z^H}{K_D + z^H} \ , \tag{4.13}$$

where κ is the number of receptors on the surface of a bacterium, K_D is the dissociation constant for binding of chemoattractants to the receptors and H is

known as the Hill coefficient. A commonly used approximation to (4.13) around the point of maximum sensitivity $\bar{z} = K_D^{1/H}$ is precisely given by

$$w(z) \approx \kappa \frac{H}{4} \ln z - \kappa \frac{\ln K_D}{4} + \frac{\kappa}{2} \ . \tag{4.14}$$

By changing the parameters κ, H and K_D, one can control α and the quality of the logarithmic approximation.

Power Laws for Stationary Densities

The expression for w in (4.11) implies the following power law for the stationary density p_∞ in terms of the attractant distribution F:

$$p_\infty(x) = c_0 F(x)^\alpha \ .$$

As α increases, the density of bacteria in the neighborhoods of the maximum of F also increases. This suggests that the factor α controls how greedy bacteria are. One can envision two main possibilities here.

In the first situation, attractants do not constitute nutrients per se, but chemical clues of where nutrients may be found. In this case bacteria seek the source that is emitting the attractants and α should be high so that bacteria find the source with high probability. This corresponds to chemotaxis working as in a search application.

In a second situation, attractants are themselves nutrients being shared by bacteria. If bacteria are greedy in this case, each bacterium will have less nutrients than if they "cooperate". Cooperation here is achieved by selecting $\alpha = 1$ so that there is an equal share of nutrients per bacterium. From Section 4.1.2, we have, for $\alpha = 1$, that the functional

$$\mathrm{E}_{p_t}\left[\ln\left(\frac{1}{2} + \frac{F}{2p_t}\right)\right]$$

increases monotonically in time to its maximum value 0. *In words, under chemotaxis with $\alpha = 1$, the average amount of nutrients per bacterium increases monotonically in time towards a state of equal share.*

This state of equal share is known in ecology as the *ideal free distribution* [50]. It is also argued in [50] that the ideal free distribution may correspond to a different power F^α due to phenomena such as interference between competitors. In [41], it was established that the ideal free distribution is a possible outcome of the advection-diffusion approximation model for chemotaxis.

One interesting point is that, as opposed to the typical prey-predator dynamics, the ideal free distribution is achieved as a result of an individualistic behavior (without direct interaction among the bacteria), which suggests that it arises as an evolutionary equilibrium. The idea of cooperation among bacteria is corroborated by the phenomenon of chemotactic signaling, according to which bacteria cooperate by emitting attractants or repellents to indicate to others the presence or scarcity of nutrients respectively [13].

In summary, we have concluded that their tumble and run behavior give bacteria the flexibility to alternate between search and deployment objectives.

Power Laws in Experiments

Only with recent advances in microfluidic devices it was possible to submit bacteria to a constant field F and measure their response with a good precision [51, 1]. This attractant profile is kept constant not only because the microfluidic devices but also because it is non-metabolizable. In these papers, chemotaxis occurs in a two-dimensional environment but with a one-dimensional gradient.

We can plot the steady-state concentrations of attractants and bacteria obtained in these papers to uncover the predicted power law. Although consistently linear for each log-log plot, the results in Figure 4.5 suggest that α may vary according to the average attractant concentrations.

A different strain of *E. coli* was considered in [1]. The power laws for the experiments in this paper reveal a more consistent slope α, see Figure 4.6. However, the linear fit with the data points given by the asterisk is not a good fit because the gradient utilized had strong discontinuities, which is not covered by

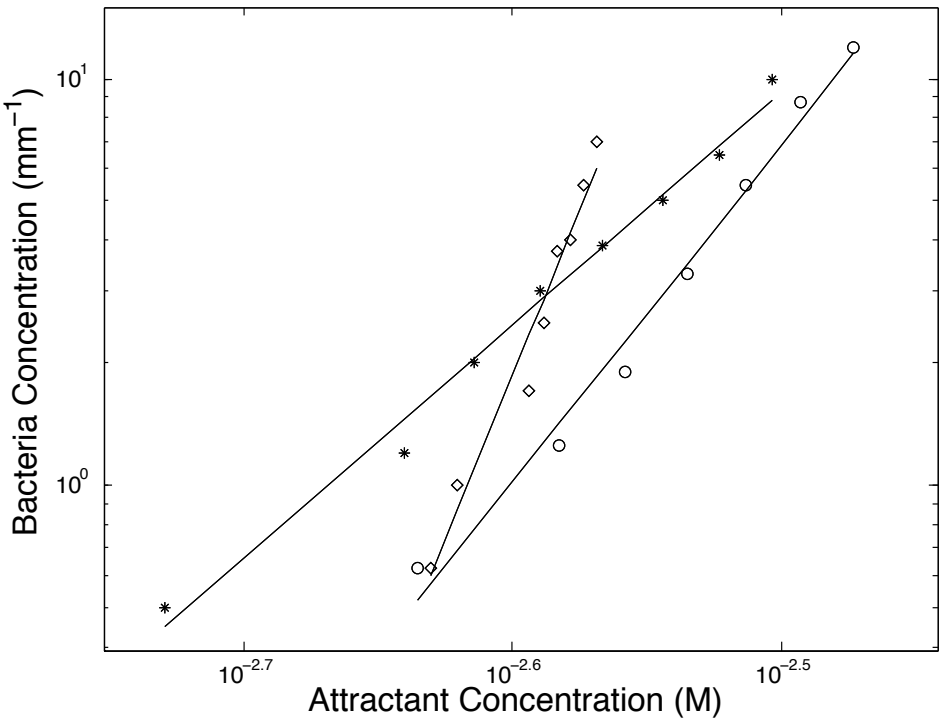

Figure 4.5: Log-log plots of bacteria concentration for the attractant gradients shown in Figures 2 and 5 of [51]
Slopes of the linear fits: $\alpha = 5.7$ (*), $\alpha = 8.3$ (circle), $\alpha = 16.1$ (diamond).

our model. Bacteria seem to be tricked to form sharp density peaks around such discontinuities.

We can also calculate α from the data analysis in these papers. Chemotaxis is modeled in these papers using the following diffusion approximation for the density p of bacteria:

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}\left(\mu_0 \frac{\partial p}{\partial x} - V_C p\right) \tag{4.15}$$

$$V_C = \frac{8V}{3\pi} \tanh\left[\frac{\chi_0 \pi}{8V}\frac{d}{dx}\left(\frac{F}{K_D + F}\right)\right] \quad , \tag{4.16}$$

where μ_0 is a diffusion coefficient, V_C represents the drift in velocity toward higher attractant concentrations, V is the two-dimensional swimming speed and χ_0 mea-

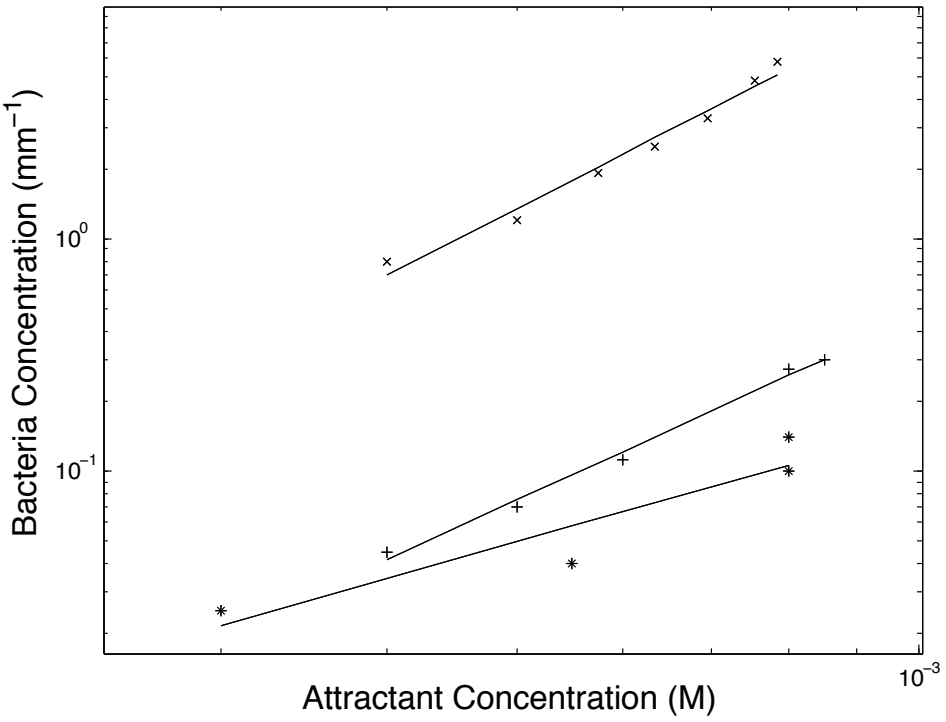

Figure 4.6: Log-log plots of bacteria concentration for the attractant gradients shown in Figures 1, 5 and 6 of [1]
Slopes of the best linear fits: $\alpha = 2.9$ (cross), $\alpha = 2.6$ (+), $\alpha = 1.6$ (*).

sures the strength of attraction of the cells to a given attractant. We can use the logarithmic approximation in (4.14) to rewrite the velocity drift as

$$V_C \approx \frac{\chi_0}{12} \frac{d \ln F}{dx} \ . \tag{4.17}$$

Plugging this approximation for V_C in (4.15) and solving for the invariant density, we obtain:

$$p_\infty \approx F^{\frac{\chi_0}{12\mu_0}} \ . \tag{4.18}$$

In [51], constant gradients of attractant ∇F were applied to bacteria and the velocity drift V_C was measured as a function of $\nabla \ln F$. In terms of the exponent

α, (4.17) and (4.18) give the relation

$$\frac{V_C}{\mu_0} = \alpha \frac{d \ln F}{dx} \ .$$

Thus, α is the slope of the curve V_C versus $\nabla \ln F$. From Figure 6 in [51], we can extract values for α in the range $[10, 30]$ with 14 being the best fit.

A different strain of *E. coli* was considered in [1], where different shapes of attractant $F(x)$ were applied to a population of bacteria and then (4.15)-(4.16) was fitted to the data by identifying μ_0 and χ_0. We can use the average values $\mu_0 = 6.6 \pm 0.6 \text{ cm}^2 \text{ s}^{-1}$ and $\chi_0 = 4.5 \pm 0.3 \text{ cm}^2 \text{ s}^{-1}$ to obtain $\alpha = 5.7 \pm 0.9$. These relatively high values of n indicate that chemotaxis works as a source-seeking mechanism for the attractant used in the experiments (α-methyl-aspartate).

Interestingly, the values of α obtained by logarithmic fitting and by model fitting seem to be off by a factor of 2 in both papers. Since χ_0 and μ_0 were calculated from the diffusion model to fit the experimental results, one should expect that they would predict a value for α consistent with the experiments. This suggests that the factor of 2 may come from some mistake in the model or in the numerical calculations. A second problem is the variation of the observed exponent α with different experiments. It is not clear whether these inconsistencies come from the model or from experiments. On one side, [51] presents exponents α that vary from experiment to experiment. On the other side, [1] presents a relatively consistent value for α. To clarify these issues, it would be valuable to perform experiments designed with the objective of measuring α.

A limitation of our model that could affect results is that we do not take into account the time that bacteria spend tumbling or the possible biases in the distribution of reorientations.

4.2.5 Bacterial Robots

A recent paper [49] has proven the idea of using bacteria as micro-actuators to be a plausible reality. In this paper, microstructures are used to attach hundreds of

bacteria to a plate in a way that their flagella are free to rotate. As a result, the plate could be moved using bacterial power. A limitation of the experiment is that it is not possible to control the motion of the plate. If one is able to tether bacteria to the plate in such a way that they are free to reorient during tumbles, then we show next that the plate will have a translational motion similar to chemotaxis. With such an artifact, it would then be possible to use chemoattractants to move the plate to some desired position.

Let N be the total number of bacteria attached to the plate and let m_i denote the propulsive force of the i-th bacterium. As in [49], the total velocity of the plate is given by

$$v = \frac{1}{k_T} \sum_{i=1}^{N} m_i \ ,$$

where k_T is the viscous drag coefficient. As in the previous sections, assume the existence of a time-invariant attractant profile $F = q^{1/n}$. Assume further that the variation of F is insignificant along the plate so that the plate can be viewed as point with position x. Using the tumbling rate model of the previous section, the tumbling rate for each bacterium is now given by

$$\lambda_i = \eta - v \cdot \nabla_x \ln q \ .$$

Let $p(x, m_1, \ldots, m_N, t)$ denote the probability density of the plate being at x with forces m_i for the i-th bacterium at time t. Then, we have the following Fokker-Planck-Kolmogorov equation for p:

$$\frac{\partial p}{\partial t} + v \cdot \nabla_x p = -\sum_{i=1}^{N} \lambda_i p + \sum_{i=1}^{N} \int_{\mathbb{M}} \lambda_i(x, m_{iC}, m') p(x, m_{iC}, m') \nu(dm') \ ,$$

where (m_{iC}, m') denotes the coordinates $(m_1, \ldots, m_N)$ with the i-th entry replaced by m'. Replacing p by q, we verify that this is indeed a stationary density:

$$\begin{aligned} 0 + v \cdot \nabla_x q &= -N(\eta - v \cdot \nabla_x \ln q) q + q \sum_{i=1}^{N} (\eta - (v - m_i/k_T) \cdot \nabla_x \ln q) \\ &= -N(\eta - v \cdot \nabla_x \ln q) q + q(N-1)(\eta - v \cdot \nabla_x \ln q) + \eta q \\ &= v \cdot \nabla_x q \ . \end{aligned}$$

Now, the same arguments as in Chapter 3 can be used to conclude that the resulting process is aperiodic Harris recurrent with invariant density q. This result means that the chemical gradient achieves the effect of synchronizing the bacteria and make them move the plate towards the same gradient. More surprisingly, perhaps, simulations show that the same convergence speed as in Figure 4.2 is obtained independently of N.

4.3 Comments and Open Problems

Besides doing simulated annealing and finding stopping rules for the search and monitoring problem, one of the most interesting extensions of the work in this chapter would be to explore the communication among agents. This could lead to faster convergence, better stopping rules or to new deployment configurations.

Exploring communication is particularly challenging if agents cannot measure their positions. In this case, the information conveyed by a neighboring agent may be its measurement of the scalar field and its relative distance or position (inferred from the radio signal). A plausible way of exploring this extra information is to add it to the function q in a sensible way and to execute the same algorithm (much like in the potential field methods in [46]).

Other relevant direction is to study the case of a time-varying scalar field q, which could represent either a moving source or the diffusion of the measured quantity.

Regarding the relationship between jump control of probability densities and chemotaxis, there are some experimental observations that need clarification and it would be useful to perform experiments directly designed for this purpose.

Chapter 5

Lyapunov Analysis

Markov processes with jumps, such as piecewise-deterministic Markov processes, offer a significant challenge to the construction of stability proofs due to the intricate nonlocal interactions in the state space that are introduced by jumps. Lyapunov stability criteria for such processes involve solving partial integro-differential inequalities, which is typically difficult to do numerically. This chapter shows that Lyapunov functions for such processes may be obtained from the minimization of a convex functional that arises in the theory of large deviations [27, 56]. In particular, we show that the Lyapunov function that maximizes a specific measure of convergence often used in large deviations theory satisfies a nonlinear integral equation without differential terms in the unknown. This equation is considerably simpler than integro-differential inequalities and we present cases in which it can be solved in closed form.

Our method was developed in the attempt to find Lyapunov functions for the PDPs in Chapters 3 and 4, where stability is proven without the use of Lyapunov techniques. The key advantage in using the Lyapunov-based techniques to prove stability of a Markov process is that one obtains information about the rate of convergence of the process to the steady-state. In particular, the method used in this chapter to construct Lyapunov functions provides conditions under which the law of the process in optimotaxis converges exponentially fast to the steady-state distribution. Information about the speed of convergence is important to

(1) estimate how long one needs to wait to be sufficiently close to the steady-state distribution and (2) design processes with fast convergence rates. This second point is exploited in Section 5.2.

The results in this chapter have the limitation that, even for a process for which the law of the process converges exponentially to the steady-state distribution, it could happen that the solution to our optimization is not a Lyapunov function. In practice, this means that one needs to verify if the function obtained by the procedure we propose is indeed a Lyapunov function, which is often a simple procedure.

5.1 Lyapunov Functions for Optimotaxis

In this section we give a Lyapunov proof for exponential ergodicity of the PDP for optimotaxis. The process represents vehicles moving with position $\mathbf{x} \in \mathsf{X} = \mathbb{R}^d$ and velocity $\mathbf{v} \in \mathbb{M} = \mathbb{S}^d$. The measure ν is the Lebesgue measure on the sphere modulo a normalization factor. The vector fields are given by $f(x, v) = v$. The jump intensity is chosen such that $\lambda(x, v) = \eta - v \cdot \nabla \ln q(x)$, where the constant η is a design parameter that must be chosen such that λ is nonnegative. As before, for such a η to exist it is necessary that $\ln q$ be globally Lipschitz. The jump kernel is such that x does not change and v has a jump distribution that is uniform on $\mathbb{M}$. More precisely,

$$Qh(x, v) = \int_{\mathbb{M}} h(x, v)\nu(dv)$$

for $h \in B(\mathcal{Y})$. We have shown in the previous chapters that a process with these characteristics has indeed an invariant density $q(x)$.

Finding Lyapunov functions for this process is difficult due to the intricate relationship between the continuous state x and the discrete mode v. To illustrate this, we consider the Metropolis-Hastings algorithm, which is a classic MCMC algorithm. Optimotaxis and Metropolis-Hastings are similar in the sense that the probabilities to reject a point in Metropolis-Hastings and the probability to

reject a velocity v in optimotaxis are essentially the same. The main difference is that, because in Metropolis-Hastings the state represents a variable in a computer, the controller can look at a point and reject it without moving the state to that point, which in turn is not possible if the state represents the position of a physical vehicle. In [77], it is shown that $q^{-1/2}$ is a Lyapunov function for the Metropolis-Hastings algorithm in terms of the goal distribution q [indeed, this is also a common Lyapunov function for diffusions]. However, it is is easy to see that no function that is independent of v can satisfy the drift condition (D1) for optimotaxis.

Using the method developed in the following sections, we were able to find a Lyapunov function for optimotaxis which turns out to be a nontrivial modification of the Lyapunov function for the random walk generated by the Metropolis-Hastings algorithm. This Lyapunov function is $u = \lambda^{1/2} q^{-1/2}$. With such a u we conclude exponential ergodicity of the PDP in the following theorem. For some $\epsilon > 0$, let η be a constant such that

$$\|\nabla_x \ln q(x)\| + \epsilon \leq \eta < \infty \ . \tag{5.1}$$

We note that a constant η is not necessary for our result, but it will simplify our proof. The next assumption characterizes distributions with exponentially decaying tails.

Assumption 6. *1.* $\|\nabla_x \ln q\|$ *is bounded*

2. $\liminf_{\|x\|\to\infty} \|\nabla_x \ln q\| > 0$

3. The Hessian $H_{xx} \ln q$ *converges to* 0 *as* $\|x\| \to \infty$.

Theorem 14. *Suppose that the output function q satisfies Assumption 6. Then, for η as in (5.1), $u = \sqrt{\lambda/q}$ is a Lyapunov function for the PDP and the PDP is exponentially ergodic:*

$$\|P^t((x,v),\cdot) - \mu\|_u \leq B_0 u e^{-b_0 t}$$

for some positive constants B_0 and b_0 and any initial condition $(x, v) \in \mathsf{X} \times \mathbb{M}$, where $d\mu = qdm$.

In the next sections, we describe the process that led to the construction of this Lyapunov function, but for now we prove this result using Theorem 5.

Proof. Because $\lambda > \epsilon$ and v is restarted uniformly after jumps, we have that, for any $A,\ C \in \mathcal{B}$ such that $\ell(A) > 0$ and C is compact, there exists a time $T < \infty$ such that the probability of reaching A from C is positive for $t \geq T$. This shows that the process is $\ell \times \nu$-irreducible and aperiodic with compact sets as petite sets. From (5.6) in the next subsection we have

$$\mathcal{L}u \leq -\frac{c_0}{2}u + b1_C \ ,$$

for $u = q^{-1/2}\sqrt{\lambda}$, a compact set C and positive constants b and c_0. We can then apply Theorem 5 to conclude the result. □

The condition of q having an exponentially decaying tail in Theorem 14 is necessary. To see why, note that the vehicles' finite velocity 1 gives a bound on how fast the support of a distribution can grow. For a vehicle starting at the origin, the stationary density is approached with an error not smaller than

$$\int_{\|x\|>t} q(x)\ell(dx)$$

at time t, which [in the symmetric scalar case] gives the exponential rate of decay

$$\lim_{r\to\infty} \frac{q(r)}{\int_{x>r} q(x)\ell(dx)} = -\lim_{r\to\infty} d(\ln q(r))/dr \ .$$

Therefore, for convergence to be exponential, q must decay exponentially. In the next subsection, we describe how we apply our method to construct a Lyapunov function for optimotaxis.

5.1.1 Constructing a Lyapunov function

Our starting point is the candidate Lyapunov function

$$u = \gamma(x)\sqrt{\lambda} \ , \tag{5.2}$$

where $\gamma(x)$ is some uniformly positive function to be identified. How we came to this candidate Lyapunov function is the theme of the next sections.

Since the considered PDP is $\ell \times \nu$-irreducible and compact sets are petite, we only need to analyze the behavior of $\mathcal{L}u$ as $\|x\|$ goes to infinity. From the definition of the generator in (2.7), we have

$$\frac{\mathcal{L}u}{u} = \frac{1}{2} v \cdot \nabla_x \ln \lambda + v \cdot \nabla_x \ln \gamma - \lambda + \sqrt{\lambda} \int \sqrt{\lambda} \, d\nu \ . \tag{5.3}$$

Define the auxiliary funcitons $\alpha := v \cdot \nabla_x \ln q$ and $\beta := \int \sqrt{\lambda} \, d\nu = \int \sqrt{\eta - \alpha} \, d\nu$. We can rewrite (5.3) as

$$\frac{\mathcal{L}u}{u} = -\frac{1}{2} \frac{v' \, \mathrm{H}_{xx} \ln q \, v}{\lambda} + v \cdot \nabla_x \ln \gamma + \alpha + \beta\sqrt{\eta - \alpha} - \eta \ ,$$

where $'$ denotes the transpose. Let $\gamma(x) = q(x)^{-k}$ for some constant $k > 0$. Then, we can rewrite

$$\frac{\mathcal{L}u}{u} = -\frac{1}{2} \frac{v' \, \mathrm{H}_{xx} \ln q \, v}{\lambda} - k\alpha + \alpha + \beta\sqrt{\eta - \alpha} - \eta \ .$$

We split the right-hand side into two parts and analyze them separately:

$$\begin{aligned} A &:= -k\alpha + \alpha + \beta\sqrt{\eta - \alpha} - \eta \\ B &:= -\frac{1}{2} \frac{v' \, \mathrm{H}_{xx} \ln q \, v}{\lambda} \ . \end{aligned}$$

Maximizing A on $\alpha \in [-\eta, \eta]$, we have the worst-case bound

$$A \leq -\eta k + \frac{\beta^2}{4\,(1-k)} \ . \tag{5.4}$$

We can find the roots of the right-hand side of (5.4) as a function of k to conclude that $A \leq 0$ for

$$\frac{1 - \sqrt{1 - \beta^2/\eta}}{2} \leq k \leq \frac{1 + \sqrt{1 - \beta^2/\eta}}{2} \ .$$

In special, $A \leq 0$ holds independently of β if and only if $k = 1/2$. When $\gamma = q^{-1/2}$, we have

$$A \leq \frac{\beta^2 - \eta}{2} \leq 0 \ , \tag{5.5}$$

where the last inequality follows from Jensen's inequality. In addition, equality holds if and only if λ does not depend on v. This analysis provides the valuable intuition that the term A in the convergence rate is taking into account how inhomogeneous the jump rate is in v or how large $\|\nabla_x \ln q\|$ is. Thus, for $\gamma = q^{-1/2}$, we have $A \leq 0$ with equality if and only if $\nabla_x \ln q = 0$.

One can also prove that the bound on A is minimized for η as small as possible. Thus, an important design principle that follows is that η must be chosen as small as possible in order to minimize the bound on A and therefore maximize the convergence rate.

To analyze the interplay between A and B, we make a distinction between two typical cases: a) when q has an exponential tail, e.g., $q = \exp(-c\|x\|)$; and b) when q has a polynomial tail, e.g., $q = \|x\|^{-c}$ for $\|x\|$ large.

Invariant density with exponential tail

In this case, $\liminf_{\|x\|\to\infty} \|\nabla_x \ln q\| > 0$ and, therefore, there is a positive constant c_0 such that $\beta^2 < \eta - c_0$ for $\|x\|$ large. On the other hand, $H_{xx} \ln q$ is bounded by a constant times $\|x\|^{-1}$ for $\|x\|$ large. Thus, A dominates B and we can use the bound in (5.5) to conclude

$$\limsup_{\|x\|\to\infty} \frac{\mathcal{L}u}{u} \leq -c_0/2 \ , \tag{5.6}$$

for $u = q^{-1/2}\sqrt{\lambda}$.

Invariant density with polynomial tail

Both A and B decay as $\|x\|^{-2}$ in this case. As a consequence, our candidate Lyapunov function cannot be used to prove exponential ergodicity. Yet, it may be

used to prove W-ergodicity. Results here depend on the specific invariant density q. Because $\beta \to 0$ as $\|x\| \to \infty$, a Lyapunov function $u = q^{-k}\sqrt{\lambda}$ maintains A nonpositive for x large only if $k = 2$. Thus, if we are interested in using this u in a Lyapunov stability proof when q has a tail of order $\|x\|^{-c}$, we need $c \geq 4$ since $\mathcal{L}u$ is of the order $\|x\|^{c/2-2}$. This is consistent with the fact that q is not a valid probability density when $c \leq 1$.

5.2 Consequences for the design

The analysis in the previous section suggests how we can change the algorithm of optimotaxis to improve convergence properties. We do this by increasing the nominal velocity. To this purpose, we redefine the vector field and the jump rate to be $f = v\rho$, for some scalar nominal velocity ρ. We can rearrange (3.13) to write the resulting jump rate as

$$\lambda = \rho(\eta - v \cdot \nabla_x \ln q\rho) \ . \tag{5.7}$$

All other parameters are kept the same.

When ρ is a constant, we have that our previous Lyapunov function $u = \sqrt{(\eta - v \cdot \nabla_x \ln q)/q}$ has its convergence rate $\mathcal{L}u/u$ in (5.3) multiplied by ρ. Provided the original process is exponentially ergodic, one can improve the convergence rate in the tail of u arbitrarily by increasing ρ. This, however, could even worsen the convergence rate to steady-state since it only takes into account the behavior on the tails. Following our analysis in the previous section, the term A would become more negative, but the term B would become necessarily more positive on the neighborhoods of the maxima of q. Intuitively, to have good convergence one wants vehicles to move slowly in the neighborhoods of the maxima of q and to escape quickly from the regions where q is small. This motivates the use of a modulated nominal velocity $\rho(q(x))$ with the properties just mentioned.

For $f = v\rho(x)$, one can verify that (5.7) still satisfies (3.13) in Theorem 12 as long as $\eta > \|\nabla_x \ln \rho q\|$. Because the divergence of f is no longer zero, we must

alter the implementation rule (3.11) to

$$\rho(\mathbf{x}(t))\mathbf{z}(t) \leq re^{\eta(t-\tau_k)}\rho(\mathbf{x}(\tau_k))\mathbf{z}(\tau_k), \;\; t \geq \tau_k \;\; .$$

For this process, one can use the method in the next sections with $p = 1/\rho$ to find the Lyapunov function $u = \sqrt{\lambda q/\rho}$ that proves exponential ergodicity provided $\|\nabla_x \ln q\rho\|$ is uniformly positive for $\|x\|$ large. This new design is illustrated in Fig. 5.1, where we see that an output dependent ρ in the range $[23, 40]$ improves speed of convergence with respect to $\rho = 25$ while ultimately providing better convergence than $\rho = 50$ [after vehicles approach the maxima of q].

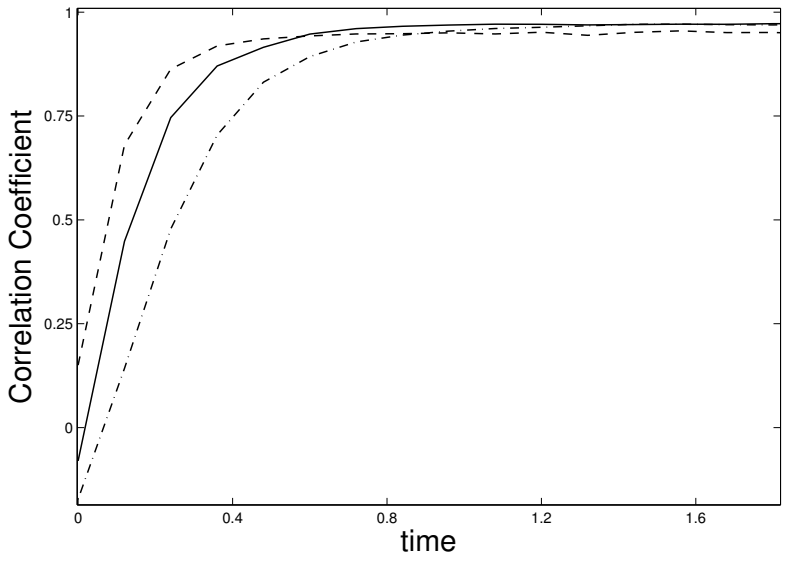

Figure 5.1: Evolution of the coefficient of correlation for an adaptive velocity. Measure function is $q(x) = 0.4e^{-\|x\|} + 0.6e^{-\|x-[1.5\ -1.5]'\|}$ for $\rho = 25$ (dot-dashed), $\rho = 50$ (dashed) and $\rho = 40\tanh([23 + 1/q^2]/40)$ (solid). Further simulation details in Chapter 4.

5.3 Lyapunov Functions and Rate Functions

In this section we discuss how one can obtain Lyapunov functions for PDPs by solving relatively simple convex optimization problems. Such Lyapunov functions are related with a notion of convergence rate that appears in the theory of large deviations of Markov processes [27]. For a probability measure ζ on $\mathcal{B}$, we define the *rate function*

$$I(\zeta) := \sup \left\{ \int -\frac{\mathcal{L}u}{u} \, d\zeta : u \in D(\mathcal{L}), u \geq 1 \right\} \ . \tag{5.8}$$

Intuitively, we can think of $-\frac{\mathcal{L}u}{u}$ evaluated at $y \in \mathsf{Y}$ as the rate of convergence of the function u at the point y and, therefore, $I(\zeta)$ would correspond to the fastest ζ-weighted average rate of convergence achievable for some function u in the domain of the generator. Rate functions have a fundamental role in the study of the probability of rare events in the context of large deviations theory. However, it is not common to solve the maximization posed in (5.8) explicitly. Our objective is to construct a Lyapunov function u by solving this maximization problem.

A converse result is given in the following proposition, which is proven in Section 5.3.4. We say that a function W_0 grows strictly slower than W if

$$\limsup_{\|x\| \to \infty} \frac{|W_0(x)|}{|W(x)|} = 0.$$

We denote this relation by $W_0 \ll W$. Consider the following drift condition:

(D3) For a constant $c > 0$, a function $W > 1$ and a small function s, the function $V : \mathsf{Y} \to [1, \infty)$ satisfies

$$\mathcal{L}V \leq -cWV + s \ .$$

Proposition 9. *Suppose that the V-exponentially ergodic process $\mathbf{\Phi}(t)$ satisfies the drift condition (D3) with W unbounded off petite sets. Then, given $u_0 \in D(\mathcal{L})$ satisfying $1 \leq u_0 \leq V$ and the growth condition $-u_0^{-1}\mathcal{L}u_0 \ll W$, there exists a unique probability measure ζ such that u_0 attains the supremum in (5.8). In*

particular, this is true for all Lyapunov functions u_0 satisfying the growth condition $-u_0^{-1}\mathcal{L}u_0 \ll W$.

In view of [56], we can rewrite the rate function in terms of an optimization problem that will be shown to be convex:

$$I(\zeta) = \sup\left\{\langle \zeta, -e^{-U}\mathcal{L}e^{U}\rangle : e^{U} \in D(\mathcal{L})\right\} \tag{5.9}$$

where $\langle\cdot,\cdot\rangle$ denotes integration. From this point on it is convenient to assume that $D(\mathcal{L})$ is an algebra and that $\psi \in D(\mathcal{L})$ implies $e^{\psi} \in D(\mathcal{L})$.

Along with convexity, another important property of the functional $\langle \zeta, -e^{-U}\mathcal{L}e^{U}\rangle$ is that it depends affinely of the differential operator $f\cdot\nabla$. As a consequence, this operator does not appear in the first variation optimality condition (5.10) or in the second variation of $\langle \zeta, -e^{-U}\mathcal{L}e^{U}\rangle$. This allows one to exploit the compactness properties typically present in jump kernels.

Some useful facts about $I(\zeta)$ are established in [56] for the discrete-time case. From [56, Prop. 4.9], we have that $I(\zeta) \geq 0$ and $I(\zeta) = 0$ if and only if ζ is an invariant measure for $\mathbf{\Phi}$, and in this case the supremum is attained by any constant function. Also from [56, Prop. 4.6] we have that $I(\zeta) < \infty$ only if ζ is absolutely continuous with respect to the invariant measure for $\mathbf{\Phi}$.

In the following theorem we provide a sufficient condition for a function to attain the supremum in (5.9). Define $K = \lambda Q$.

Theorem 15. *i. The optimization problem in (5.9) is a convex optimization problem in U.*

ii. A sufficient condition for $u = e^{U} \in D(\mathcal{L})$, $U \geq 0$, to attain the supremum in (5.8) for $d\zeta = p\ d(\ell\times\nu)$, $p \in D(\mathcal{L}^*)$, is

$$uK^*\left(\frac{p}{u}\right) - \nabla\cdot fp - \frac{p}{u}Ku = 0 \quad \ell\times\nu\text{-a.e.} \tag{5.10}$$

iii. Suppose e^{U} is a solution to (5.10). Then, $e^{G} \in D(\mathcal{L})$ is also a solution if and only if $G(y) - U(y) = G(z) - U(z)$ $\zeta(dz)K(z,dy)$-a.e.

The proof is provided in Section 5.3.3. Roughly speaking, Theorem 15 *(iii)* implies that the set of solutions to (5.10) is invariant under multiplication by harmonic functions of Q (i.e., functions h such that $Qh = h$). In particular, this is always true for multiplication by a constant.

5.3.1 Computation of Optimizers

The main result in this section lies in the observation that solving (5.10) for u can be done with relative ease for a significant number of PDPs. This is true because (5.10) has no differential terms in u, which makes it possible to exploit the compactness properties of K. In particular, when K and K^* have finite rank, solving (5.10) reduces to a finite-dimensional problem. A wide class that satisfies this property is given by a generalization of the Markov Jump Linear systems in [31], where one may allow Markov transitions to depend on the continuous state. This is also the case of our optimotaxis example, where a closed form solution to (5.10) is provided.

To see why this is possible, note that we can rewrite (5.10) using the fact that u solves an implicit quadratic equation:

$$u = \frac{\nabla \cdot pf + \sqrt{(\nabla \cdot pf)^2 + 4pK(u)K^*(p/u)}}{2K^*(p/u)} \quad . \tag{5.11}$$

When K and K^* have finite rank, this expression defines a finite dimensional manifold where u lies. Therefore, even when it is not possible to solve (5.11) explicitly, this equation gives us structure to make good guesses for candidate Lyapunov functions.

One can also attempt to solve (5.11) via an iterative procedure as follows: given u_n, replace u on the the right hand side of (5.11) with u_n and define u_{n+1} to be that value modulo some normalization. The normalization is necessary since the class of solutions to (5.10) is invariant under multiplication by a constant. Under reasonable conditions, this iteration is verified to converge for finite rank jump kernels.

A possible but more difficult technique to solve the convex functional minimization is to use gradient methods as presented in [97] for example. The success of these methods depends on the gradient being locally Lipschitz and uniformly monotone, which are properties that seem to be linked with exponential ergodicity.

Finally, it is important to remark that this is the point where our approach takes advantage of the specialized setting of PDPs. If, for example, one was to consider a purely deterministic process, the solution set to (5.10) would be trivial (either empty or the whole space of functions) and u cannot be used as a Lyapunov function. On the other hand, nontrivial results can be obtained in the nondeterministic case. But, when one considers a general process that includes both jumps and diffusion, the solution to (5.10) is typically difficult, since one would be dealing with a partial integro-differential equation.

5.3.2 A Candidate Lyapunov Function for Optimotaxis

Although $K = \lambda Q$ is not a compact operator in $\mathcal{B}(\mathsf{Y} = \mathsf{X} \times \mathbb{M})$, it is a finite rank operator in $\mathcal{B}(\mathbb{M})$ for every fixed $x \in \mathsf{X}$. In fact, if we regard Q as an operator in $\mathcal{B}(\mathbb{M})$ for a fixed x, its range is spanned by the constant function. Moreover, the operator Q has the property that $Q(\alpha(x)\psi) = \alpha(x)Q\psi$ for any $\psi \in B(\mathsf{Y})$ and any α independent of v. This implies that the set of solutions to (5.10) is invariant under multiplication by a function of x only.

We obtain from (5.11)

$$u = \frac{v \cdot \nabla p + \sqrt{(v \cdot \nabla p)^2 + 4\lambda p \int u \, d\nu \int \frac{\lambda p}{u} \, d\nu}}{2 \int \frac{\lambda p}{u} \, d\nu} \ .$$

This implies that there exist functions r and s such that u satisfies the following structure

$$u = r(x) \left(v \cdot \nabla p + \sqrt{(v \cdot \nabla p)^2 + \lambda p s(x)} \right) \ .$$

Let p be a multivariable normal distribution and let its covariance tend to infinity. The u that results from the limit is equivalent to that we would obtain

with $p = 1$, but in this case p would not be integrable. Although our theory has no need to restrict ζ to be a probability measure, we avoid this path due to its more complicated interpretation. The resulting limit satisfies

$$u = r(x)\sqrt{\lambda s(x)} \ .$$

Recalling that the set of solutions to (5.10) is invariant under multiplication by a function of x only, we have

$$u = \gamma(x)\sqrt{\lambda} \ . \tag{5.12}$$

for any $\gamma(x)$ such that $u \in D(\mathcal{L})$.

This u can be interpreted as the function that maximizes the rate of convergence with equal weight for every (x, v). It is clear that not all elements of the form (5.12) are Lyapunov functions for the PDP. However, we have arrived to a structure for Lyapunov functions without which we were not able to find Lyapunov functions in the past.

5.3.3 Proof of Theorem 15

General convexity results analogous to the ones in the theorem have a straightforward derivation in the discrete-time case as in [56]. Unfortunately, we have to be more careful in the continuous-time case and explicitly use the generator of the process. Define

$$\mathcal{H}(G) := e^{-G}\mathcal{L}e^{G} = f \cdot \nabla G + \lambda(e^{-G}Qe^{G} - 1) \ .$$

To verify convexity of $\langle \zeta, \mathcal{H}(\cdot)\rangle$, we note that, for constants $\alpha, \beta > 0$ and $F, G \in D(\mathcal{L})$,

$$\langle \zeta, \mathcal{H}(\alpha F + \beta G)\rangle = \alpha\langle \zeta, f \cdot \nabla F\rangle + \beta\langle \zeta, f \cdot \nabla G\rangle + \int \zeta(dz)K(z, dy)e^{\alpha(F(y)-F(z))+\beta(G(y)-G(z))} \ .$$

Convexity therefore follows from the convexity of the exponential function. Moreover, the convexity inequality is strict unless

$$F(y) - G(y) = F(z) - G(z) \;\; \zeta(dz)K(z, dy)\text{-a.e.}$$

This proves item (i). To prove (ii), we consider the difference

$$\mathcal{H}(G + F) - \mathcal{H}(G) = f \cdot \nabla F + \lambda e^{-G} Q e^{G} (e^{-F} Q_g e^{F} - 1) \;\; ,$$

where $Q_g h = \frac{Q(e^G h)}{Qe^G}$ is a Markov kernel. By a convexity argument, we have that

$$e^{-F} Q_g e^{F} - 1 \geq \ln e^{-F} Q_g e^{F} \geq Q_g F - F \;\; .$$

This allows us to write

$$\mathcal{H}(G + F) - \mathcal{H}(G) \geq f \cdot \nabla F + \lambda e^{-G} Q e^{G} (Q_g F - F) =: \mathcal{L}_g F \;\; . \tag{5.13}$$

Following [56], for example, $\mathcal{L}_g$ is known as the generator of the twisted semigroup and it can be written in the general form

$$\mathcal{L}_g h = e^{-G} \mathcal{L}(e^{G} h) - h e^{-G} \mathcal{L} e^{G}$$

for $g = e^G$. Conditions under which the generator of twisted semigroup is the Fréchet derivative of $\mathcal{H}$ are given in [56]. Here, to keep our set of assumptions minimal, we use $\mathcal{L}_g$ only as a subdifferential of $\mathcal{H}$. Therefore, a sufficient condition for U to be a maximizer of $\langle \zeta, -\mathcal{H}(\cdot) \rangle$ is that ζ be an invariant measure for the twisted semigroup generated by $\mathcal{L}_u$. Indeed, in this case we have

$$\langle \zeta, \mathcal{H}(U + F) - \mathcal{H}(U) \rangle \geq \langle \zeta, \mathcal{L}_u F \rangle = 0 \;\; ,$$

for all $F \in D(\mathcal{L})$. A sufficient condition for ζ, $d\zeta = p \; d(\ell \times \nu)$, to be an invariant measure under $\mathcal{L}_u$ is that $\mathcal{L}_u^* p = 0$, which can be written as (5.10). Finally, we note that necessary and sufficient conditions could be obtained following the more general framework in [56].

5.3.4 Proof of Proposition 9

From our characterization of optimizers in the proof of Theorem 15, it is sufficient to find ζ such that it is the invariant probability measure of the semigroup generated by $\mathcal{L}_{u_0}$. Next we show that the $\mathcal{L}_{u_0}$ satisfies the drift condition (D2) and, therefore, the measure ζ exists uniquely. From (5.13), we have that the process generated by $\mathcal{L}_{u_0}$ has the same drift term as the original process and a jump term such that, if a jump is probable for the original process, it is also probable for the new process. This implies that the process generated by $\mathcal{L}_{u_0}$ inherits irreducibility from the original process. This process satisfies the drift condition with Lyapunov function $V_0 = V/u_0$:

$$\mathcal{L}_{u_0} V_0 = \frac{\mathcal{L}(u_0 V_0) - V_0 \mathcal{L} u_0}{u_0} = V_0 (V^{-1} \mathcal{L} V - u_0^{-1} \mathcal{L} u_0) \ .$$

The growth condition $u_0^{-1} \mathcal{L} u_0 \ll W$ and the fact that W is unbounded off petite sets guarantee that there exist constants $c_0 > 0$, $b_0 < \infty$ and a petite set C_0 such that

$$\mathcal{L}_{u_0} V_0 \leq -c_0 W V_0 + b_0 1_{C_0} \ .$$

5.4 Example: A Scalar Linear Transport Process

In this section we show how the method proposed above can be applied to a process different from optimotaxis. The process we consider in this example appears in many fields ranging from neutron transport phenomena to biology (see [73] and references therein). This well-studied process provides a simple example of how our method can be used to find Lyapunov functions in closed form. This process describes a particle with position $\mathbf{x} \in \mathsf{X} = \mathbb{R}$ moving with velocity $\mathbf{v} \in \mathbb{M} = \{-1, +1\}$. Velocity jumps may occur with intensity $\lambda_v(x) \geq \epsilon > 0$ and with jump kernel $Q = \delta_{\{-v\}}$, where δ denotes the Dirac mass. A process so defined is aperiodic and irreducible, with compact sets being petite.

From (5.11), we can derive an optimizer of the form

$$u_v(x) = \frac{v \cdot \nabla p_v + \sqrt{(v \cdot \nabla p_v)^2 + 4p_v \lambda_v u_{-v} \lambda_{-v} p_{-v}/u_{-v}}}{2p_{-v}\lambda_{-v}/u_{-v}} \ .$$

As above, take the limit as $p \to 1$. The resulting limit satisfies

$$u_v(x) = u_{-v}\sqrt{\frac{\lambda_v}{\lambda_{-v}}} \ .$$

Therefore, the class of minimizers is characterized by functions of the form

$$u_v(x) = \gamma(x)\sqrt{\lambda_v}$$

where the function γ depends only on x . To construct a Lyapunov function, we need now to select γ properly. To this purpose, we evaluate

$$\frac{\mathcal{L}u_v}{u_v} = v(\ln\gamma)' + \frac{v}{2}(\ln\lambda_v)' + \sqrt{\lambda_v\lambda_{-v}} - \lambda_v$$

where $'$ denotes derivative with respect to x. In order to make the right-hand side negative for both values of v, we must choose γ so that

$$-\frac{1}{2}(\ln\lambda_{-1})' + \sqrt{\lambda_1\lambda_{-1}} - \lambda_{-1} \ < \ (\ln\gamma)' \ < \ -\frac{1}{2}(\ln\lambda_1)' - \sqrt{\lambda_1\lambda_{-1}} + \lambda_1 \ .$$

Assuming that the inequalities hold, we choose $(\ln\gamma)'$ to be the mean of the two bounds:

$$(\ln\gamma)' = -\frac{1}{2}(\ln\sqrt{\lambda_1\lambda_{-1}})' + \frac{1}{2}\lambda_1 - \frac{1}{2}\lambda_{-1} \ .$$

This leads to a consistent choice of $\gamma \geq 1$ as long as $\lambda_1 \geq \lambda_{-1}$ in the positive tail and $\lambda_1 \leq \lambda_{-1}$ in the negative tail. For this choice of γ we have

$$\frac{\mathcal{L}u_v}{u_v} = \frac{1}{2}\left(\ln\frac{\sqrt{\lambda_1}}{\sqrt{\lambda_{-1}}}\right)' + \sqrt{\lambda_1\lambda_{-1}} - \frac{\lambda_1+\lambda_{-1}}{2} \ . \tag{5.14}$$

For x large enough, the derivative term in the right-hand side of (5.14) is either dominated by λ_1 or negative. Therefore, exponential ergodicity depends solely on the difference between the geometric and the arithmetic averages of the jump

intensities, which is uniformly negative provided $|\lambda_1 - \lambda_{-1}| > 0$ uniformly outside a compact set.

In summary, we have constructed a Lyapunov function that predicts exponential ergodicity for the transport process given that there exists $\epsilon > 0$ such that $\mathrm{sgn}(x)(\lambda_1 - \lambda_{-1}) > \epsilon$ for x large enough. If this condition holds with ϵ depending on x but decaying more slowly than $1/|x|$, one can show that $\mathcal{L}u_v < 0$ uniformly off a compact set, which is the drift condition (D1) for ergodicity. These results are consistent with the expression for the stationary probability density $q_v(x)$ for this process [73]:

$$q_v(x) + q_{-v}(x) = c \exp\left(\int_0^x (\lambda_{-1} - \lambda_1)\, dz \right) \quad . \tag{5.15}$$

This density is integrable if and only if $\mathrm{sgn}(x)(\lambda_1 - \lambda_{-1}) > 0$ decays more slowly than $1/|x|$, which shows that our candidate Lyapunov function provides nonconservative conditions for ergodicity. Although this is not the focus of our contribution, it is worth mentioning that, to the best of our knowledge, there is no similar Lyapunov approach to this problem in the literature. Finally we note that, despite the similarities with optimotaxis, this example shows that our technique can construct Lyapunov functions for processes with jump kernel qualitatively different.

5.5 Comments and Open Problems

We have presented a method for the construction of Lyapunov functions for PDPs based on the maximization of a certain notion of rate of convergence. This method allowed us to construct a Lyapunov function to prove exponential ergodicity for the optimotaxis algorithm. On the optimotaxis end, it would be interesting to use Lyapunov functions to design annealing schemes and to predict convergence rates. Lyapunov functions also provide a norm in which the process behaves continuously with variations in the model and, therefore, can be used to predict robustness properties of the process.

Regarding our approach to construct Lyapunov functions, some open questions are how one can select ζ such that the maximizer function is guaranteed to be a Lyapunov function; and to find general conditions under which the iterative procedure suggested in Section 5.3.1 converges to the solution of the optimization problem.

Chapter 6

Redundant Data Transmission in Control/Estimation over Lossy Networks

In the new generations of wireless communications (3G and 4G), channel adaptive techniques are able to provide large improvements in almost any performance metric. These techniques utilize an adaptive allocation of communication resources as the channel conditions change with time. This chapter explores a similar idea in the context of networked control systems (NCS). In addition to compensating for the uncertainty generated by the channel, we aim at allocating communication resources to compensate for the uncertainty being generated within the controlled process. These adaptive techniques are well-suited for NCSs because they permit an increase in the reliability of communication without increasing the transmission delays, to which NCSs have low tolerance.

Adaptation can be achieved by adjusting the transmit power, by adaptive coding (an example of which is changing the quantization coarseness) and by *diversity* schemes, which consist of the transmission of redundant signals through mostly independent channel realizations. Diversity schemes may involve using multiple time slots, frequency slots, antennas or network paths [86]. While many diversity schemes are dynamically exploited in data networks by scheduling transmissions according to the network status (see e.g. [30, 94]), these techniques do not take

into account nor benefit from the dynamical nature of NCSs. In this chapter we focus on diversity schemes and show that something as simple as transmitting multiple independent copies of the same packet can provide significant performance gains in the context of NCSs.

The adopted NCS architecture is depicted in Fig. 6.1, which considers the case of a single sensor and a controller. We assume that, by means of some diversity scheme, a number of independent redundant channels is available for data transmission. These are erasure channels with i.i.d. probabilities of packet drop out. At each time step, the sensor sends measurements to the controller with a certain level of redundancy. Further, by means of an acknowledgement mechanism, the sensor knows which measurements were received by the controller. Our focus is on deciding how many redundant copies of a packet should be transmitted at each sampling time and what benefits can be drawn from this.

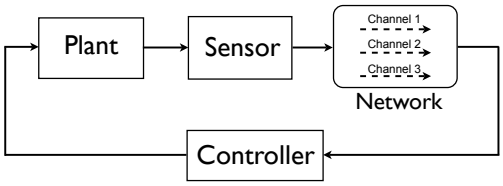

Figure 6.1: NCS architecture

The basic intuition behind this technique is that one can use redundancy to increase the probability of a successful transmission whenever the estimation error in the controller becomes large. On the other hand, at time instants for which the control performance is satisfactory, one may send only one packet or not send data at all, which would save communication resources. This adaptive behavior is desirable for NCSs because it improves the reliability of transmissions without relying on error correction schemes that induce delay in the transmissions. Indeed, if a packet containing some measurement data is dropped at a time instant, it is generally more important from a control point of view to guarantee that the

measurements at the next time instant are delivered, rather than to retransmit the old information that was dropped previously.

For simplicity of presentation, we consider NCSs with full local state measurements and no network delays. However, the results obtained can be readily extended to the case of partial state measurements and delays in the network as suggested by Xu in [92]. In the case of partial measurements, it is ideal to use a Kalman filter collocated with the sensor and then transmit optimal estimates.

In the first part of the chapter we focus our attention on the problem of stabilizing a discrete-time linear time-invariant process with a certainty equivalence control. For such process it is well know that mean-square instability arises whenever the drop probability rises above a certain threshold. Moreover, no matter how small the drop probability is, some statistical moments of the process state will always be unbounded. It turns out that redundant transmissions can be used to stabilize any given statistical moments for any probability of drop. Surprisingly, we show that, by a judicious use of redundant transmissions, this can be achieved with no significant increase in the average communication rate.

In the second part, motivated by the observation that the redundant channels are rarely used and yet their availability provides significant performance gains, we discuss the possibility of having multiple nodes sharing the redundant channels. This is a natural practice to maximize network efficiency but one must be aware that transmissions at the different feedback loops become interdependent. Nevertheless, we provide protocols for which mean-square stability is preserved for an arbitrary number of nodes sharing the same set of channels.

In the third part, we investigate optimal redundant transmission protocols for the NCS. In this setting, the controller constructs estimates of the process state using the measurements transmitted by the sensor. The sensor, in turn, uses a redundant transmissions policy that minimizes the weighted average cost of the estimation error in the controller and the average communication rate. First, we consider the ideal case in which the sensor has enough computational power to

reconstruct the state estimates available to the controller from the acknowledgement information. Secondly, and motivated by the fact that in some applications sensors have very limited computational capabilities, we find policies that minimize the same cost but base their decisions solely on the number of consecutive transmission failures.

6.1 A linear NCS with redundant transmissions

Throughout the chapter we consider a linear time-invariant plant with

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + B\mathbf{u}(k) + \mathbf{w}(k) \tag{6.1}$$

where $\mathbf{x} \in \mathbb{R}^n$ denotes the state of the process, $\mathbf{u} \in \mathbb{R}^{n_1}$ the control input, and $\mathbf{w} \in \mathbb{R}^n$ an n-dimensional zero-mean Gaussian white noise process with positive definite covariance matrix Σ. The pair (A, B) is assumed to be controllable.

The controller and a sensor that measures the full state $\mathbf{x}(k)$ are connected through a network that drops packets independently of each other, with probability $p \in (0, 1)$. The state $\mathbf{x}(k)$ is assumed to be transmitted with negligible quantization error. In order to "adjust" the probability that the measurement $\mathbf{x}(k)$ reaches the controller, the sensor may transmit multiple copies of this message through independent channels (see Fig. 6.1). The transmitter is equipped with a feedback channel that allows it to know which packets are dropped. We denote by $\mathbf{l}(k)$ the number of consecutive transmission failures that occurred immediately before time k, where a *transmission failure* is characterized by the failure of all the attempts to transmit $\mathbf{x}(k)$ at time k.

We are interested in designing protocols that determine how many identical packets to send at time k as a function of $\mathbf{x}(k)$ and $\mathbf{l}(k)$. We denote by $\mathbf{v}(k)$ the number of redundant packets transmitted at time k. Under our assumption of independent drops, the probability of a transmission failure at time k is $p^{\mathbf{v}(k)}$.

We adopt a certainty equivalence control law of the form

$$\mathbf{u}(k) = K\hat{\mathbf{x}}(k) \tag{6.2}$$

where the matrix K is chosen such that $A + BK$ is Schur and $\hat{\mathbf{x}}(k)$ is an estimate of $\mathbf{x}(k)$ based on the measurements that successfully reached the controller up to time k. In particular,

$$\hat{\mathbf{x}}(k) := \mathrm{E}\left[\mathbf{x}(k)\middle|\mathbf{x}(s),\ s < k,\ s \in \mathcal{T}_{\text{success}}\right] \tag{6.3}$$

where $\mathcal{T}_{\text{success}}$ denotes the set of times at which the sensor succeeded in transmitting the measured state to the controller. This state estimate can be computed recursively using

$$\hat{\mathbf{x}}(k+1) = \begin{cases} A\hat{\mathbf{x}}(k) + B\mathbf{u}(k) & \text{if } k \notin \mathcal{T}_{\text{success}} \\ A\mathbf{x}(k) + B\mathbf{u}(k) & \text{if } k \in \mathcal{T}_{\text{success}}. \end{cases} \tag{6.4}$$

Subtracting (6.1) from (6.4), we conclude that the estimation error $\mathbf{e}(k) := \hat{\mathbf{x}}(k) - \mathbf{x}(k)$ evolves according to the dynamics:

$$\mathbf{e}(k+1) = \begin{cases} A\mathbf{e}(k) - \mathbf{w}(k) & \text{if } k \notin \mathcal{T}_{\text{success}} \\ -\mathbf{w}(k) & \text{if } k \in \mathcal{T}_{\text{success}}. \end{cases} \tag{6.5}$$

The closed-loop dynamics (6.1)–(6.2) can be expressed in terms of this error using

$$\mathbf{x}(k+1) = (A + BK)\mathbf{x}(k) + BK\mathbf{e}(k) + \mathbf{w}(k). \tag{6.6}$$

In the following sections we investigate stability and optimal estimation in this setting. From (6.6), our certainty equivalence control guarantees a bounded covariance for $\mathbf{x}(k)$ if $\mathbf{e}(k)$ has bounded covariance.

6.2 Moment Stabilization Using Redundant Transmissions

In this section we investigate the stability properties of redundant transmission protocols that can be specified by a static law v that maps the number $\mathbf{l}(k)$ of

consecutive transmission failures to the number $v(\mathbf{l}(k))$ of packets to send. For example, if we use the identity function $v(l) = l$, then $\mathbf{l}(k)$ identical packets will be sent at time k. The possibility of sending zero packets is not excluded.

Theorem 16 shows that mean-square stability can be achieved for any system matrix A and any drop probability $p < 1$ by a suitable choice of the redundant packet transmission protocol that specifies the function $v(l)$.

Theorem 16. *Let the spectral radius of the matrix A be denoted by a. The covariance of $\mathbf{x}(k)$ is bounded when the following limit exists and satisfies*

$$\lim_{l\to\infty} a^2 p^{v(l)} < 1 \ .$$

Conversely, the covariance of $\mathbf{x}(k)$ is unbounded when $\lim_{l\to\infty} a^2 p^{v(l)} > 1$.

Therefore, all that is needed to guarantee stability is to select $v(l)$ sufficiently large for large values of l:

$$\lim_{l\to\infty} v(l) > \frac{2\log a}{-\log p} \ . \tag{6.7}$$

Proof. We assume that $a \geq 1$ since the covariance is always bounded if $a < 1$. In view of (6.6) and the fact that $A + BK$ in (6.6) is Schur, it is sufficient to verify the boundedness of the covariance matrix for $\mathbf{e}(k)$. To this purpose, we consider the infinite Markov chain for $\mathbf{l}(k)$, which, under the assumption of independent drops, has transition probabilities

$$\begin{aligned} &\Pr\left(\mathbf{l}(k+1) = \mathbf{l}(k) + 1 \mid \mathbf{l}(k)\right) = p^{\mathbf{l}((k))}, && k \geq 0 \\ &\Pr\left(\mathbf{l}(k+1) = 0 \mid \mathbf{l}(k)\right) = 1 - p^{v(\mathbf{l}(k))}, && k \geq 0 \ . \end{aligned}$$

The stationary probabilities $\mu(l)$ for this Markov chain must therefore satisfy

$$\begin{aligned} &\mu(l+1) = p^{v(l)}\mu(l) = p^{\sum_{m=0}^{l} v(m)}\mu(0), \ l \geq 0 \\ &\sum_{l=0}^{\infty} \mu(l) = 1 \ , \end{aligned}$$

which allows us to conclude that

$$\mu(0) = \left(1 + \sum_{l=1}^{\infty} p^{\sum_{m=0}^{l-1} v(m)}\right)^{-1}$$

and, for $l > 1$,

$$\mu(l) = p^{\sum_{m=0}^{l-1} v(m)} \left(1 + \sum_{m=1}^{\infty} p^{\sum_{n=0}^{m-1} v(n)}\right)^{-1} . \tag{6.8}$$

Notice that $\mu(0)$ is well defined since, by (6.7) and the fact that $a \geq 1$, there exits a constant $L > 0$ such that $v(l) \geq 1,\ \forall l \geq L$. Under this condition one can also verify that the chain is aperiodic and recurrent. Therefore, we can apply [69, Thm. 14.3.3] to conclude that

$$\lim_{k\to\infty} \mathrm{E}\left[\mathbf{e}(k)\mathbf{e}(k)'\right] = \sum_{l=0}^{\infty} \mathrm{E}\left[\mathbf{e}(k)\mathbf{e}(k)' \,\middle|\, \mathbf{l}(k) = l\right] \mu(l) \ . \tag{6.9}$$

In view of (6.5), we have that

$$\begin{aligned}\mathrm{E}\left[\mathbf{e}(k)\mathbf{e}(k)' \,\middle|\, \mathbf{l}(k)\right] &= \mathrm{E}\Big[\left(\sum_{m=0}^{l} A^m \mathbf{w}(k-m)\right) \cdot \left(\sum_{m=0}^{l} A^m \mathbf{w}(k-m)\right)'\Big] \\ &= \sum_{m=0}^{l} A^m \Sigma A'^m \ ,\end{aligned}$$

which can be used in (6.9) to obtain

$$\lim_{k\to\infty} \mathrm{E}\left[\mathbf{e}(k)\mathbf{e}(k)'\right] = \sum_{l=0}^{\infty} \mu(l) \sum_{m=0}^{l} (A^m \Sigma A'^m) \ . \tag{6.10}$$

For any submultiplicative matrix norm $\|\cdot\|$, we have

$$\| \lim_{k\to\infty} \mathrm{E}\left[\mathbf{e}(k)\mathbf{e}(k)'\right] \| \leq \|\Sigma\| \sum_{l=0}^{\infty} \mu(l) \sum_{m=0}^{l} (\|A^m\|^2) \ .$$

By the ratio test, this series is convergent if

$$\lim_{l\to\infty} \frac{\mu(l+1)}{\mu(l)} \frac{\|A^{l+1}\|^2}{\|A^l\|^2} = a^2 \lim_{l\to\infty} p^{v(l)} < 1 \ ,$$

where the equality comes from (6.8) and the fact that $a{=}\lim_{k\to\infty} \|A^k\|^{1/k}$. This gives the first part of the theorem. The second part can be deduced by pre- and post-multiplying (6.10) by the eigenvector corresponding to a and then using the ratio test to conclude divergence of the resulting series. □

For the stability of higher moments, one can obtain conditions analogous to (6.7) with a similar proof.

From (6.7), we can see that to achieve stability one may require a protocol that, at times, sends a large number of packets, which seems to require a large communication rate. To verify that this is not the case, we investigate the expected communication rate for a given function $v(l)$. We assume that the packet size is a constant with value 1 and that it is sufficiently large so that the controller receives $\mathbf{x}(k)$ with negligible quantization loss. We define the expected asymptotic transmission rate as

$$\bar{R} := \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathrm{E}[v(\mathbf{l}(k))] \ . \tag{6.11}$$

Theorem 17. *Suppose that $a^2 p^M < 1$ for some integer M. Then, for every integer $N \geq 0$, there exists a protocol with $v(l) \leq M$ that stabilizes the covariance of $\mathbf{x}(k)$ with an expected transmission rate:*

$$\bar{R} = O(1/N) \ , \tag{6.12}$$

which can be made arbitrarily small by choosing N sufficiently large.

Proof. Consider the protocol

$$v(l) = \begin{cases} 0 & \text{for } l \leq N \\ M & \text{for } l > N \ . \end{cases} \tag{6.13}$$

From Theorem 16, this is a stabilizing protocol. As in the proof of Theorem 16, we can use (6.8) and the recurrence of the chain to compute the expected asymptotic transmission rate

$$\bar{R} = \mu(0)\left(v(0) + \sum_{l=1}^{\infty} v(l) p^{\sum_{m=0}^{l-1} v(m)} \right) \ .$$

Substituting (6.13), we have that

$$\bar{R} = \frac{M}{N(1-p^M)+1} \ ,$$

which gives the result in the theorem. □

While we can obtain an arbitrarily small communication rate, the larger we make N the larger the error covariance will be. This relationship between average transmission rate and control performance is investigated in the following sections.

We note that a strong result as (6.12) may no longer hold when quantization errors cannot be neglected, because packet sizes could need to increase when many consecutive failures take place to compensate the quantization errors as $\mathbf{x}(k)$ gets large.

A second factor that may alter our results is the presence of failures in the acknowledgement mechanism. While there may be many ways to deal with this issue such as using redundancy in the acknowledgement channel, one can always adopt the conservative approach of using positive acknowledgements. This approach preserves stability since it sends necessarily more redundant packets than in the case of perfect acknowledgement.

6.3 Redundant Protocols with Multiple Nodes

As seen in the previous section, redundant protocols may need to utilize redundant channels quite rarely. This suggests that some of these channels could be shared by other feedback loops without compromising performance. In this section we explore the situation in which multiple processes share the redundant channels.

Consider the case where S feedback loops share the redundant channel. Denote by N the total number of channels, by p the dropout probability for each independent channel. For this scenario, we define the following protocol.

Protocol 1. *A node i does not use redundancy as long as its number $\mathbf{l}_i(k)$ of consecutive drops is not larger than a constant L; when $\mathbf{l}_i(k) > L$ for node i only, that node uses all the N available channels; if $\mathbf{l}_i(k) > L$ for more than one node, these nodes share the N channels in an arbitrary way.*

This protocol does not require a fair share of the N channels, but only that the channels are utilized without drops due to collision. This would work either in a centralized scenario or in a decentralized scenario where transmitters take advantage of the *capture effect* [99] so that, even if multiple nodes access the same channel, one of them is still able to successfully use it.

We are interested in evaluating under which conditions Protocol 1 stabilizes the estimation error in the mean-square sense. Somewhat surprisingly, we show next that stability can be achieved independently of the number of processes S in the network (even for a possibly decentralized protocol). Let A_i be the system matrix of the i-th process and a_i its spectral radius. Then we define $a = \max_{1 \le i \le S}\{a_i\}$.

Theorem 18. *Suppose that $a^2 p^N < 1$. Then, for any number of feedback loops S, there always exists an integer L such that Protocol 1 stabilizes the estimation error in the mean-square sense.*

Proof. Let $V(l) := \max_{j \in S} l_j$. To prove mean-square stability, it suffices to prove the boundedness of $E[W(\mathbf{l}(k))]$ for $k \ge 0$, where the Lyapunov function $W(l)$ is defined as $W(l) = a^{2V(l)}$. To this purpose, we now evaluate $\Delta W(l_0) := E[W(\mathbf{l}(L)) - W(\mathbf{l}(0))|\mathbf{l}(0) = l_0]$.

When $V(l_0) < L$, we have $V(l(L)) < 2L$ with probability 1. This gives

$$\Delta W(l_0) \le -W(l_0) + a^{4L}, \text{ for } V(l_0) < L \ . \tag{6.14}$$

When $V(l_0) \ge L$, we have that $V(l(L)) < 2L$ if at least S successful transmissions occur in the first L time steps. Otherwise, $V(l(L))$ may be as large as $V(l_0) + L$. Let P_L be the maximum probability that less than S nodes transmit

in L time steps. This probability is maximum when when $\mathbf{l}_j(0) \geq L$ for all $j \leq S$ and when only one node transmits in all channels. Then, for $S < L$, we have

$$\begin{aligned} P_L &= \sum_{i<S} \left(\begin{array}{c} L \\ i \end{array} \right) (1-p^N)^i \; p^{N(L-i)} \\ &= (L-S+1) \Big(\begin{array}{c} L \\ S-1 \end{array} \Big) \int_0^{p^N} t^{L-S}(1-t)^{S-1} \; dt \\ &\leq \frac{L-S+1}{L-S} \Big(\begin{array}{c} L \\ S-1 \end{array} \Big) p^{N(L-S+1)} \;\; , \end{aligned} \tag{6.15}$$

where the second equality comes from the cumulative distribution of binomial random variables and the inequality comes from the bound $(1-t) \leq 1$. With this, we conclude that

$$\begin{aligned} \Delta W(l_0) &\leq P_L a^{2(V(l_0)+L)} + 1 \cdot a^{4L} - a^{2V(l_0)} \\ &= (a^{2L} P_L - 1) W(l_0) + a^{4L} \end{aligned} \tag{6.16}$$

for $V(l_0) \geq L$.

Because the right-hand side of (6.16) dominates the right-hand side of (6.14), we have that (6.16) holds for all l_0. Therefore, by a standard Lyapunov argument, $E[W(\mathbf{l}(kL))]$ is bounded provided that

$$a^2 P_L^{1/L} < 1 \;\; . \tag{6.17}$$

Since the boundedness of $E[W(\mathbf{l}(kL))]$ does not depend on the initial condition, it also implies the boundedness of $E[W(\mathbf{l}(k))]$. Finally, from (6.15), note that

$$\begin{aligned} \lim_{L\to\infty} P_L^{\frac{1}{L}} &\leq \lim_{L\to\infty} \left(\frac{L-S+1}{L-S} \right)^{\frac{1}{L}} \Big(\begin{array}{c} L \\ S-1 \end{array} \Big)^{\frac{1}{L}} p^{N(L-S+1)/L} \\ &= p^N \;\; . \end{aligned}$$

Since $a^2 p^N < 1$, one can always choose L large enough to satisfy (6.17). □

For scenarios in which capacity is degraded by collisions, a stability result will depend highly on the specifics of the communication system. As in (6.17) in the proof of Theorem 18, stability will depend on the quantity $\lim_{L\to\infty} P_L^{1/L}$, which is closely related to the notion of error exponent in information theory [24], i.e., the logarithm of the probability of error in a block of length L.

6.4 Optimal Communication Protocols

In the last two sections, we have seen that it is possible to stabilize one or several processes in the mean-square sense, with very few communication resources. However, this may lead to large error covariances. Our goal now is to determine an optimal policy for the single node case that decides when to send multiple copies of the same packet and how many copies to send. This policy should be optimal in the sense that it achieves an optimal trade off between the conflicting objectives of keeping small the estimation error $\mathbf{e}(k)$ that drives the closed-loop dynamics (6.6) while achieving this with a minimal amount of communication.

To formulate this problem, we adopt the framework of Markov Decision processes on Borel spaces of Section 2.4. The state is the estimation error $\mathbf{e} \in \mathsf{X} = \mathbb{R}^d$ and the control action is the number of redundant packets $\mathbf{v} \in \mathsf{B} \subset \mathbb{N}$. We denote by $B(\mathbf{e})$ of admissible control actions. The probability transition kernel can be extracted from (6.5) as

$$P(dy|e,v) = (1-p^v)f(y)dy + p^v f(y-Ae)dy \ , \tag{6.18}$$

where f is the pdf of the normal distribution with zero mean and covariance Σ.

For technical reasons, we restrict the set of admissible control actions B so that all policies under consideration are stabilizing in a uniform way (as it will be clear later). In particular,

$$B(e) := \begin{cases} \{0,\dots,M\} & \text{if } \|e\| < L \\ \{M\} & \text{if } \|e\| \geq L \ , \end{cases} \tag{6.19}$$

where M denotes the maximum number of redundant packets possible and $L > 0$ is a constant. This restriction ensures that M packets are sent when $\|e\|$ grows as large as L, but our results allow L to be arbitrarily large.

Our objective is the minimization of the following average cost (AC) criterion

$$J(\pi, e_0) := J_{\text{est}}(\pi, e_0) + \lambda J_{\text{com}}(\pi, e_0) \tag{6.20}$$

where

$$J_{\text{est}}(\pi, e_0) := \lim_{N\to\infty} \frac{1}{N} \mathrm{E}_{e_0}^{\pi} \Big[\sum_{k=0}^{N-1} \mathbf{e}(k)' Q \mathbf{e}(k) \Big] \tag{6.21}$$

$$J_{\text{com}}(\pi, e_0) := \lim_{N\to\infty} \frac{1}{N} \mathrm{E}_{e_0}^{\pi} \Big[\sum_{k=0}^{N-1} \mathbf{v}(k) \Big] \tag{6.22}$$

where λ is a positive scalar, Q a positive definite matrix and $\mathrm{E}_{e_0}^{\pi}$ denotes the expectation given a policy π and an initial state $\mathbf{e}(0) = e_0$.

The criterion in (6.20) is a weighted sum of two terms: the first term $J_{\text{est}}(\pi, e_0)$ penalizes a time-averaged expected quadratic estimation error, whereas the second term $J_{\text{com}}(\pi, e_0)$ penalizes the average communication rate, measured in terms of the number of messages sent per unit of time. The constant λ allows one to adjust the relative weight of the two terms. As $\lambda \to 0$, communication is not penalized, whereas as $\lambda \to \infty$, communication is heavily penalized. Intermediate values of λ will yield Pareto-optimal compromise solutions between the two conflicting criteria.

Our assumption on the set of admissible control actions guarantees that the limits in the definitions of J_{est} and J_{com} exist. Hence, we can use (6.5) to rewrite the cost function (6.20) in terms of the one-step cost c as follows

$$J(\pi, e) = \lim_{N\to\infty} \frac{1}{N} \mathrm{E}_{e}^{\pi} \sum_{k=0}^{N-1} c(\mathbf{e}(k), \mathbf{v}(k)) \tag{6.23}$$

where

$$c(e, v) = e' Q e + \lambda v \ . \tag{6.24}$$

In the next theorem we apply the results from Section 2.4 to state the existence of a solution to the AC-optimality problem.

Theorem 19. *Suppose that the maximum redundancy degree M in (6.19) is sufficiently large so that*

$$a^2 p^M < 1 \ . \tag{6.25}$$

Then:

1. *There exist a triple $(\varrho^*, \phi^*, \pi^*)$ satisfying the (ACOE) (2.10).*
2. *π^* is AC-optimal and ϱ^* is the optimal AC-function.*
3. *The value iteration algorithm of Section 2.4 converges (with rate at least exponential).*

A proof of Theorem 19 is provided in Appendix A.

6.4.1 Suboptimal Protocols

Solving for the optimal policy in Theorem 19 may be computationally intense for high-dimensional systems. Fortunately, one can exploit the linear structure of the controlled plant to find suboptimal policies that are computationally tractable. One approach to construct a reasonable suboptimal policy consists of applying just one iteration of the policy iteration algorithm (see [35]) initialized with the policy $\pi_0 \equiv M$. The resulting policy is

$$\pi_1(e) = \arg \min_{v \in B(e)} \{\lambda v + p^v e' A' H A e\} \ , \tag{6.26}$$

where the positive definite matrix H satisfies the Lyapunov equation

$$p^M A' H A - H + Q = 0 \ .$$

The policy π_1 is considerably simpler to compute than π^* and it offers the additional benefit of a smaller (memory-wise) representation. Alternatively, one may

construct similar suboptimal policies by searching for a quadratic relative value function $\bar{\varphi}(e) = e'He$ that minimizes the average cost upper bound $\bar{\varrho}$:

$$\begin{aligned}\bar{\varrho} + \bar{\varphi} &\geq \min_{v \in B(e)} \left[c(e,v) + \int \bar{\varphi}(y) P(dy|e,v) \right] \\ &= \min_{v \in B(e)} \left[e'Qe + \lambda v + p^v e' A' H A e + \operatorname{tr} H\Sigma \right] \ . \end{aligned} \tag{6.27}$$

As the numerical examples in Section 6.6 show, the policy π_1 is already quite close to the optimal and further optimization may be unnecessary.

6.5 A Simplified Optimal Protocol

In many applications sensors have limited computational capabilities that could prevent the use of elaborate protocols that require the computation of estimation errors as those considered in Section 6.4. To address this issue, we can design a simplified protocol that bases its decision rule only on the consecutive number of failures $\mathbf{l}(k)$ that occurred prior to the k-th sampling time, much like the protocols considered in Section 6.2.

In general, this approach would lead to a partially observable Markov decision process. Fortunately, due to the special structure of the problem, we show that it is enough to consider a fully observable process. We consider the two controlled Markov chains $(\mathbf{e}(k), \mathbf{l}(k))$ and $\mathbf{l}(k)$ as described in the previous sections. Let Π_l denote the set of stabilizing feedback policies for which $\mathbf{v}(k)$ depends only on $\{\mathbf{l}(s); s \leq k\}$. For $\pi \in \Pi_l$ and the chain $\mathbf{l}(k)$, define the average cost

$$J_l(\pi, l_0) = \lim_{N \to \infty} \frac{1}{N} \operatorname{E}_{l_0}^{\pi} \sum_{k=1}^{\infty} \bar{c}(\mathbf{l}(k), \mathbf{v}(k)) \ , \tag{6.28}$$

where the one-step cost is given by

$$\bar{c}(l, v) = \operatorname{tr}(Q\Sigma_l) + \lambda v \ , \tag{6.29}$$

where

$$\Sigma_l := \sum_{m=0}^{l} A'^m \Sigma A^m \ . \tag{6.30}$$

Theorem 20. *For $\pi \in \Pi_l$,*

$$J(\pi, e_0) = J_l(\pi, l_0), \ \forall e_0, l_0 \ , \tag{6.31}$$

where J is the original cost in (6.20). Therefore, an AC-optimal policy for the cost J_l and the chain $\mathbf{l}(k)$ is also an AC-optimal policy for the cost J and the chain $(\mathbf{e}(k), \mathbf{l}(k))$ within the set of policies Π_l.

Proof. We can rewrite our cost criterion as

$$J(\pi, e_0) = \lim_{N\to\infty} \frac{1}{N} \mathrm{E}^{\pi}_{e_0} \sum_{k=1}^{\infty} \mathrm{E}^{\pi}_{e_0} \left[c(\mathbf{e}(k), \mathbf{v}(k)) \,\middle|\, \mathbf{l}(k) \right] \ . \tag{6.32}$$

Once a packet is successfully transmitted, the belief (conditional probability) of $\mathbf{e}(k)$ is solely given by $f(e)$, through (6.18), and it does not depend on any previous beliefs. Hence, the average cost criterion does not depend on the initial belief. Thus, without loss of generality, we can calculate the average costs assuming that $l_0 = 0$ and e_0 has distribution $f(\cdot)$. For this initial condition, $\mathbf{e}(k)$ is conditionally distributed as $\sum_{m=0}^{\mathbf{l}(k)} A^m(k)\boldsymbol{\omega}_m$ given $\mathbf{l}(k)$, where $\boldsymbol{\omega}_m$ are i.i.d. variables with density $f(\cdot)$. Then, $\mathrm{E}^{\pi}_{e_0} \left[c(\mathbf{e}(k), \mathbf{v}(k)) \,\middle|\, \mathbf{l}(k) = l \right] = \bar{c}(l, v)$ and the claim of the theorem follows from (6.32) and (6.28). □

If we further restrict the set of policies to be such that $v = p^M$ for $l \geq T$, we no longer need to keep track of the number of consecutive drops when this number exceeds T. We can truncate the Markov chain $\mathbf{l}(k)$ by redirecting the jumps $(l = T) \to (l = T + 1)$ to $(l = T) \to (l = T)$. Thus, we have moved from the infinite dimensional problem in Section 6.4 to a finite dimensional problem. Using the per-stage cost $\bar{c}$ and the transition probabilities for $\mathbf{l}$ that we described in Section 6.2, one can calculate AC-optimal policies that depend on $\mathbf{l}$ only. This could be done either via dynamic programming or via direct optimization, since the average costs can be directly calculated using the stationary distribution as in Section 6.2. Interestingly, once $\bar{c}$ is known, the complexity of solving for the optimal protocol does not depend on the dimension n of the dynamical system, but only on T and on the size $M + 1$ of the set of control actions.

In general, the designer can select the constants M and T to be small. This is so because the probability of failure p^M becomes indistinguishably small for M larger than some small constant, typically 3. Likewise, the probability of reaching a state $l = T$ also decreases exponentially with T. An important consequence of the set of control actions being small is that one can solve for the optimal protocol offline and then use small look-up tables. The same observations also apply to the optimal protocols in the previous section.

Remark 5. *When solving an optimal control problem with quadratic costs, the separation principle does not hold in the setting of Section 6.4 since the communication protocols may depend on* $\mathbf{x}(k)$*. On the other hand, the separation principle does hold in the context of the simplified protocols of this section. To see why, note that, for each fixed simplified protocol, we have a Markov Jump Linear System, for which the separation principle is known to hold [31]. As it turns out that the optimal control and optimal estimator do not depend on the fixed protocol, the separation principle holds for the general problem.*

6.6 Numerical Examples

6.7 Example 1 – A scalar process

The results in the previous sections were applied to a scalar example with $A = 2$, $\Sigma = 3$, $Q = 1$, $p = 0.15$ and $L = 10$. By varying λ from 0.001 to 200, we constructed the Pareto frontiers shown in Fig. 6.2. Some important observations can be deduced from this figure:

1. With protocols where $v \in \{1, 2\}$, we are able to decrease the estimation cost by 30% while increasing the communication cost by only 6%.

2. The simplified optimal policy discussed in Section 6.5 produces protocols that can be quite close to the Pareto-optimal boundary.

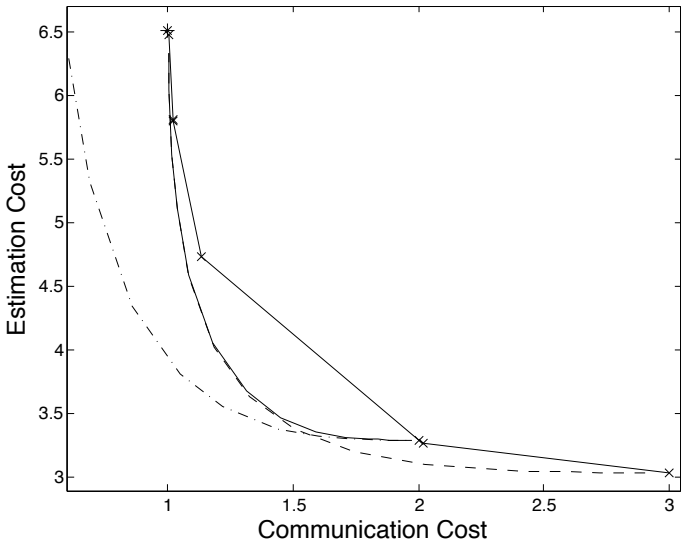

Figure 6.2: Pareto Frontiers.
Optimal policy with $v \in \{1, 2\}$ (solid); optimal policy with $v \in \{1, 2, 3\}$ (dashed); optimal policy with $v \in \{0, 1, 2\}$ (dash-dotted); simplified optimal policy with $v \in \{1, 2, 3\}$, $l < T = 5$, and $v = 3$ for $l \geq T = 5$ (cross).

3. Increasing the maximum number of redundant packets beyond 2 hardly improves the Pareto-optimal boundary.

4. Most simplified optimal policies are nontrivial in the sense that their redundancy degree is not constant (leading to non-integer communication costs).

5. If we were to allow no transmissions at some time instants (i.e., $\mathbf{v}(k) \geq 0$ instead of $\mathbf{v}(k) \geq 1$) then one could further improve the optimal Pareto-optimal boundary (dash-dotted line in Fig. 6.2).

A phenomenon that commonly arises in multi-objective MDPs is that points on the Pareto frontier do not always correspond to deterministic policies. This is the case for the Pareto frontier of the simplified protocol, where only the points marked with a cross correspond to deterministic policies and the lines linking those points correspond to randomized policies that can be derived from the deterministic ones as explained in Section 2.4.1.

Figure 6.3 illustrates the fact that the use of optimal policies becomes more advantageous as the drop out probability p is increased. As p increases, the performance of the simplified protocols are worsened significantly with respect to the optimal protocols. On the other hand, the performances of protocols with $v \in \{1, 2\}$ and $v \in \{1, 2, 3\}$ remain close to each other for the range of p shown in the figure. Interestingly, the difference between optimal costs for the optimal policy with $v \in \{0, 1, 2\}$ and that with $v \in \{1, 2\}$ remains constant with p.

6.8 Example 2 – Two dimensional process

For the following analysis, we have considered a two dimensional example with $A = [1\ 1; 0\ 1]$, $\Sigma = [0.75\ 0; 0\ 0.75]$, $Q = [1\ 0; 0\ 1]$, $p = 0.15$ and $L = 10$.

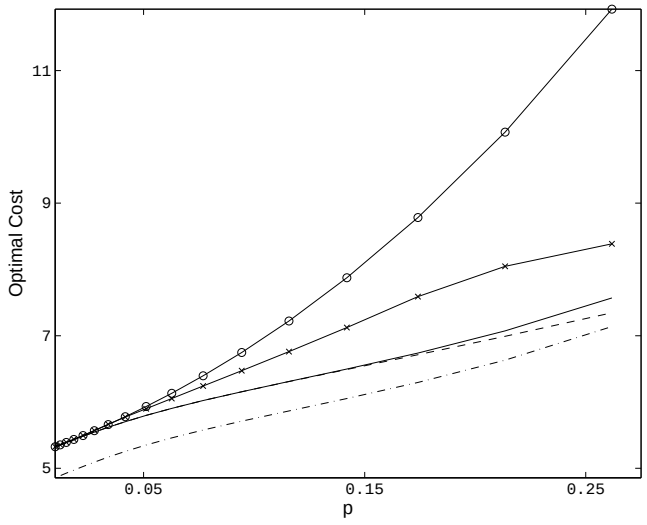

Figure 6.3: Optimal costs as a function of the drop probability.
Optimal policy with $v \in \{1, 2\}$ (solid); optimal policy with $v \in \{1, 2, 3\}$ (dashed); optimal policy with $v \in \{0, 1, 2\}$ (dash-dotted); simplified optimal policy with $v \in \{1, 2, 3\}$, $l < T = 5$, and $v = 3$ for $l \geq T = 5$ (cross); policy $v(k) \equiv 1$ (o).

6.8.1 Single node

By varying λ from 10^{-5} to 10, we constructed the Pareto frontiers shown in Fig. 6.6, where policies are restricted to the different action sets $v \in \{1, 2\}$, $v \in \{1, 2, 3\}$ and $v \in \{0, 1, 2\}$. Similar conclusions as in the previous example can be drawn.

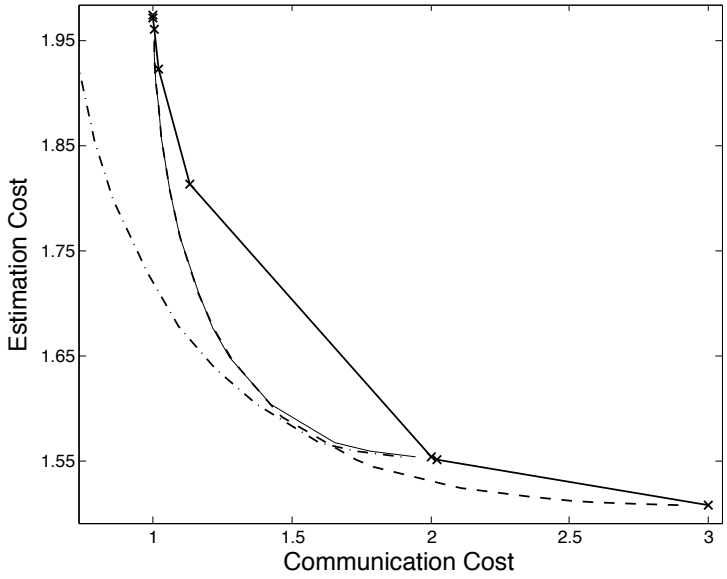

Figure 6.4: Pareto Frontiers.
Optimal policy with $v \in \{1, 2\}$ (solid); optimal policy with $v \in \{1, 2, 3\}$ (dashed); optimal policy with $v \in \{0, 1, 2\}$ (dash-dotted); simplified optimal policy with $v \in \{1, 2, 3\}$, $l < T = 5$, and $v = 3$ for $l \geq T = 5$ (cross).

Suboptimal approaches have been considered in Fig. 6.5. To show the performance improvement that arises from judiciously sending redundant information, we considered also the baseline policy that always sends one packet per time step. Some observations from this figure are:

1. The suboptimal policy π_1 in (6.26) gives a performance remarkably close to the Pareto frontier and significantly better than the simplified protocols.

2. Using the trivial policy $\mathbf{v}(k) = 1$, $\forall k$, leads to less communication (x-axis) than the policies that use $v \in \{1, 2\}$ and $v \in \{1, 2, 3\}$, but this is at the

expense of a significantly larger estimation error (y-axis). In fact, based on the results of Section 6.2, we know that for unstable systems and large drop probabilities, $\mathbf{v}(k) \equiv 1$ can lead to instability.

3. The steep slope of the Pareto frontier at communication cost 1 indicates that a large percentual decrease in estimation cost can be obtained with a small percentual increase in communication cost.

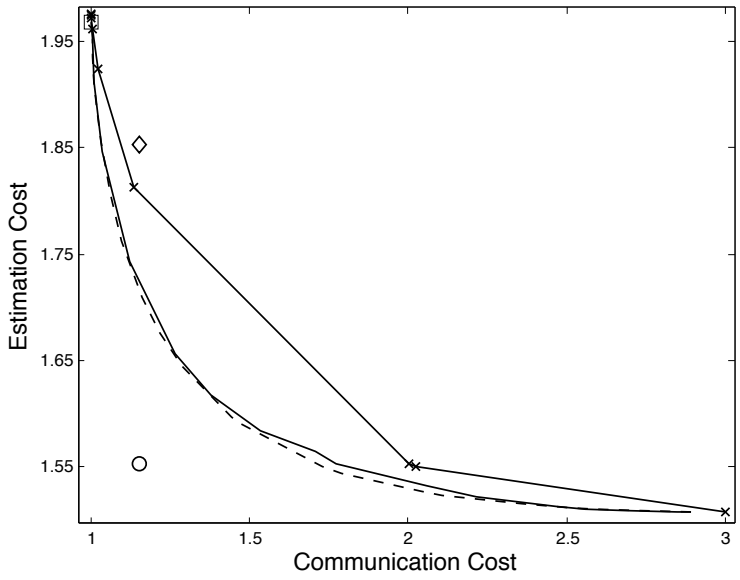

Figure 6.5: Optimal vs. suboptimal policies.
Optimal policy with $v \in \{1, 2, 3\}$ (dashed); simplified optimal policy with $v \in \{1, 2, 3\}$, $l < T = 5$, and $v = 3$ for $l \geq T = 5$ (cross); suboptimal policy π_1 with $v \in \{1, 2, 3\}$ (solid); policy $\mathbf{v}(k) = 1\ \forall k$ (square); TCP with no delay (circle); TCP with delayed retransmission (diamond).

6.8.2 Redundant Transmissions vs. Retransmissions

The motivation for using redundant transmissions as opposed to retransmissions (such as in TCP) is that NCSs performance may suffer due to the delay accumulated in the transmission, error detection, acknowledgements and retransmission.

To provide some intuition on how the protocols described here compare with TCP we consider two idealized TCP-like protocols. In the first protocol, we consider a TCP-like scheme in which it is always possible to perform one transmission of $\mathbf{x}(k)$ and, if necessary, at most one retransmission before the sampling time $k+1$. At time $k+1$, the measurement $\mathbf{x}(k+1)$ is available at the sensor for transmission and the old measurement $\mathbf{x}(k)$ is discarded (regardless of whether it was successfully transmitted). In the second scheme, a retransmission of $\mathbf{x}(k)$ is only received at time $k+1$ so it cannot be used until time $k+2$. Hence, the retransmitted message is only used if the message containing $\mathbf{x}(k+1)$ is dropped. We assume that original and retransmitted packets are dropped with i.i.d. probabilities with the same value as in the redundant channels above.

The results are shown in Fig. 6.5. The delay free TCP scheme outperforms the redundant transmission protocols for a large region of cost combinations. On the other hand, a one-time-step delay is already enough to have the TCP scheme outperformed by the redundant protocols. In practice, many things can go wrong with TCP such as random delays and correlation between drops, which in our framework would imply a higher probability that the retransmitted packets are not used (dropped). Irrespectively of these implementation issues, redundant protocols offer the advantage of flexibility in the choice of different estimation and communication costs combinations by selecting different operating points within the Pareto boundary.

6.8.3 Multiple nodes

To illustrate the possibility of sharing redundant channels among multiple nodes, we consider an example with two nodes and three channels. Each node has access

to its own channel and shares a redundant channel with the other node. Rather than utilizing Protocol 1, that was defined mainly to prove the scalability results in Section 6.3 and may not be very practical, we have both nodes utilizing the single-node simplified optimal protocol derived in the previous example. The dynamics for each node is also the same as in the previous example. Two possible collision resolution schemes are considered. In *Scheme* 1, if a collision happens on the redundant channel, both packets are dropped on that channel. In *Scheme* 2, one of the nodes is granted the use of the redundant channel with probability 1/2 despite of collisions. The performance results for these schemes are shown in Fig. 6.6. For lower average communication costs (below 1.13), these schemes perform close to the single-node case. Whereas *Scheme* 2 still provides improvement in the estimation cost as the communication cost grows, *Scheme* 1 has its estimation error performance degraded when the nodes access the redundant channel more than 13% of the time.

6.9 A Note on Dynamic Stabilization and Output Feedback

We consider now the output feedback case

$$\begin{aligned} \mathbf{x}(k+1) &= A\mathbf{x}(k) + B\mathbf{u}(k) + \mathbf{w}(k) \\ \mathbf{y}(k) &= C\mathbf{x}(k) + \mathbf{r}(k) \ , \end{aligned} \tag{6.33}$$

where $\mathbf{y}(k)$ is the output measured by the sensor and $\mathbf{r}(k)$ is the measurement noise. We assume that (A, B) is controllable and (A, C) observable. Given a controller $G(z)$ that stabilizes the system with perfect communication, we want to construct a stabilizing controller for the system with drops that produces the same output as $G(z)$ when there are no drops. To this purpose, we can use the Youla parametrization to rewrite the control input for the system without drops as

$$\bar{\mathbf{u}}(k) = K\bar{\mathbf{x}}(k) + F(q)\bar{\mathbf{e}}(k) \ , \tag{6.34}$$

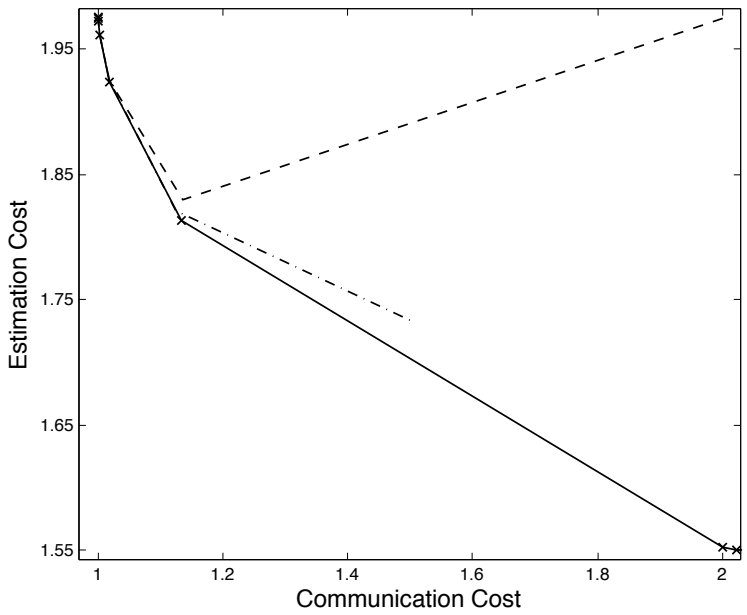

Figure 6.6: Performance in the two-node case.
Simplified policy for *Scheme* 1 with $v \in \{1, 2\}$ and $T = 5$ (dashed); simplified policy for *Scheme* 2 with $v \in \{1, 2\}$ and $T = 5$ (dash-dot). Simplified optimal policy for the single-node case with $v \in \{1, 2\}$ and $T = 5$ (cross).

where K is such that $A+BK$ is Schur, $F(q)$ is some stable transfer function, $\bar{\mathbf{x}}(k)$ is a state estimate satisfying

$$\bar{\mathbf{x}}(k+1) = (A - LC)\bar{\mathbf{x}}(k) + L\mathbf{y}(k) + B\mathbf{u}(k) \ ,$$

for some stable observer gain L, and $\bar{\mathbf{e}}(k) = \mathbf{x}(k) - \bar{\mathbf{x}}(k)$ is the estimation error and satisfies

$$\bar{\mathbf{e}}(k+1) = (A - LC)\bar{\mathbf{e}}(k) + \mathbf{w}(k) - L\mathbf{r}(k) \ .$$

For any stabilizing controller $G(z)$, there exists a corresponding stable transfer function $F(q)$ such that (6.34) is a realization of such controller. This suggests that a simple way of extending $G(z)$ to the networked control system is to have the controller to keep estimates of $\bar{\mathbf{x}}(k)$ and $\bar{\mathbf{e}}(k)$.

There are two major approaches to generate estimates of $\mathbf{x}$ in the networked case. In the first approach, the sensor constructs and transmits estimates of $\mathbf{x}(k)$. In the second approach, the sensor sends all the available measurements that were not yet received by the controller and estimates are constructed at the controller. The first approach is more computationally intensive for the sensor, but, depending on the dimensions of $\mathbf{x}$ and $\mathbf{y}$ and on the probability of drops, it may use less memory and communication resources. If the likelihood of consecutive drops is low and the dimension of $\mathbf{x}$ is large compared to $\mathbf{y}$, the second approach is preferable in general. If we adopt the first approach, we need that both $\bar{\mathbf{x}}(k)$ and $\bar{\mathbf{e}}(k)$ be transmitted in order to implement our control (6.34). With either approach, the controller can construct the estimate $(\hat{\mathbf{x}}(k), \hat{\mathbf{e}}(k))$ given by

$$\hat{\mathbf{x}}(k+1) = \begin{cases} A\hat{\mathbf{x}}(k) + LC\hat{\mathbf{e}}(k) + B\mathbf{u}(k) & \text{if } k \notin \mathcal{T}_{\text{success}} \\ A\bar{\mathbf{x}}(k) + LC\bar{\mathbf{e}}(k) + B\mathbf{u}(k) & \text{if } k \in \mathcal{T}_{\text{success}} \end{cases}$$

and

$$\hat{\mathbf{e}}(k+1) = \begin{cases} (A - LC)\hat{\mathbf{e}}(k) & \text{if } k \notin \mathcal{T}_{\text{success}} \\ (A - LC)\bar{\mathbf{e}}(k) & \text{if } k \in \mathcal{T}_{\text{success}}. \end{cases}$$

It is now possible to define a control input for the networked control system that is compatible with the controller $G(z)$:

$$\hat{\mathbf{u}}(k) = K\hat{\mathbf{x}}(k) + F(q)\hat{\mathbf{e}}(k) \ .$$

The error $\tilde{\mathbf{u}} := \bar{\mathbf{u}} - \hat{\mathbf{u}}$ in the control input is then given by

$$\tilde{\mathbf{u}}(k) = K\tilde{\mathbf{x}}(k) + F(q)\tilde{\mathbf{e}}(k) \ ,$$

where $\tilde{\mathbf{x}}(k) := \bar{\mathbf{x}}(k) - \hat{\mathbf{x}}(k)$ and $\tilde{\mathbf{e}}(k) := \bar{\mathbf{e}}(k) - \hat{\mathbf{e}}(k)$ satisfy

$$\begin{bmatrix} \tilde{\mathbf{x}}(k+1) \\ \tilde{\mathbf{e}}(k+1) \end{bmatrix} = \begin{bmatrix} A & LC \\ 0 & A-LC \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}(k) \\ \tilde{\mathbf{e}}(k) \end{bmatrix} + \begin{bmatrix} I & L \\ 0 & -L \end{bmatrix} \begin{bmatrix} \mathbf{w}(k) \\ \mathbf{r}(k) \end{bmatrix} \tag{6.35}$$

if $k \notin \mathcal{T}_{\text{success}}$, and

$$\begin{bmatrix} \tilde{\mathbf{x}}(k+1) \\ \tilde{\mathbf{e}}(k+1) \end{bmatrix} = \begin{bmatrix} I & L \\ 0 & -L \end{bmatrix} \begin{bmatrix} \mathbf{w}(k) \\ \mathbf{r}(k) \end{bmatrix}$$

if $k \in \mathcal{T}_{\text{success}}$. We can apply the results in Section 6.2 to conclude the existence of protocols that provide a bounded covariance for $[\tilde{\mathbf{x}}(k) \ \ \tilde{\mathbf{e}}(k)]'$ when

$$p^M \rho\left(\begin{bmatrix} A & LC \\ 0 & A-LC \end{bmatrix} \right) = p^M \rho(A) < 1 \ ,$$

where M is the maximum number of available channels and $\rho(\cdot)$ denotes the spectral radius of a matrix. Since $\bar{\mathbf{u}}$ is stabilizing, $\hat{\mathbf{u}}$ guarantees a bounded covariance for $\mathbf{x}$ provided that $\tilde{\mathbf{u}}$ has bounded covariance. On the other hand, since $F(q)$ is stable, the covariance of $\tilde{\mathbf{u}}$ is bounded if $[\tilde{\mathbf{x}}(k) \ \ \tilde{\mathbf{e}}(k)]'$ has bounded covariance. Therefore, the same stability condition as in the state feedback case is recovered. Moreover, by making the error covariance of $[\tilde{\mathbf{x}}(k) \ \ \tilde{\mathbf{e}}(k)]'$ small, we bring the system performance closer to that provided by the control $G(z)$.

6.10 Comments and Open Problems

We introduced new communication protocols for networked control systems that adjust the probability of successful communication by the transmission of redun-

dant packets. These results can be readily extended to the case of partial state measurements and delays in the network by following the procedure in [92].

Our results suggest that redundant channels may be efficiently shared among multiple processes. In addition, the proposed technique has a diminishing returns property in the sense that little additional benefits are obtained by increasing the number of redundant channels beyond two or three. This implies that the implementation of the diversity schemes will not demand an extremely expensive infra-structure.

As future work, it is important to address the case when the drops for different packets are not independent of each other. This would be important to study communication faults due to collisions when this type of redundancy strategy is simultaneously employed by different nodes.

One should also consider the case in which nodes do not share the same information on what was broadcasted to the network, e.g., acknowledgements are not perfect. The development of new acknowledgement mechanisms would be a valuable approach in this case. In particular, there are cases where nodes can efficiently detect the occurrence of drops through the plant (as opposed to an acknowledgement signal in the network) as described in [29].

Chapter 7

A Framework for General Capacity Scheduling

Besides diversity, one can also adjust capacity by varying the transmit power or the coding scheme, an example of which would be changing the quantization coarseness. Within diversity techniques, one could also adjust the capacity of available channels. Our next goal is to propose a framework to design general capacity scheduling protocols.

Varying the transmit power gives a somewhat simple variation of the results of the previous chapter. Indeed, the transmit power controls the bit error rate, which in turn controls the drop probability. To handle quantization, however, we need to consider more general channels than erasure channels. To this purpose, we consider the following control system:

$$\begin{aligned} \mathbf{x}(k+1) &= A\mathbf{x}(k) + B\mathbf{u}(k) + \mathbf{w}(k) \\ \mathbf{y}(k) &= C\mathbf{x}(k) + \mathbf{b}(k)\boldsymbol{\omega}(k) + \mathbf{v}(k) \ , \end{aligned}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the state, $\mathbf{y} \in \mathbb{R}^{n_y}$ is the output received by the controller, $\mathbf{u} \in \mathbb{R}^{n_u}$ and $\mathbf{b} \in \mathbb{R}$ are control inputs, $\mathbf{w} \in \mathbb{R}^n$ is the process noise, $\mathbf{v} \in \mathbb{R}^{n_y}$ is the measurement noise and $\boldsymbol{\omega} \in \mathbb{R}^{n_y}$ is the noise resulting from the network transmission. A, B and C are the system matrices with the appropriate dimensions.

In this formulation, we can regard the probability density of $\mathbf{b}(k)\boldsymbol{\omega}(k)$ as being controlled by means of varying the transmit power, the quantization or coding

scheme, or the number of repeated packets. While it may be true that controlling the noise amplitude in a multiplicative fashion is not possible depending of the way capacity is adjusted, this is a reasonable approximation for our design purposes. This model is general enough to approximate most ways of controlling capacity. For the erasure channels of the previous sections, for example, a drop would correspond to the event of the variance of $\boldsymbol{\omega}(k)$ being very large.

This framework may also incorporate different channel models. For example, a typical channel model is to consider two-state Markov chain: the noise covariance of $\boldsymbol{\omega}(k)$ is small in one state and large in the other.

A more precise model would consist of $\boldsymbol{\omega}(k)$ being a vector with different noise models that are selected by the control $\mathbf{b}(k)$. The design principles in this chapter also apply to this type of model.

In a Markov Decision Process framework, we can assign different costs to $\mathbf{b}(k)$ depending on the method used to control capacity. In addition, depending on the encoding being used, it might make more sense to consider that the signal to noise ratio is being controlled instead of the absolute noise. This would be the case for example if logarithmic quantization is being used. To cope with this case, we can make the communication cost depend on the error covariance $\mathbf{P}$ of the estimator.

Another important variation that can be captured by this model is the case of estimation at the sensor. For all practical purposes, we can set C to be the identity and to incorporate the estimation error in $\mathbf{v}(k)$.

7.1 Protocol Design

We assume that the controller constructs estimates of the state $\mathbf{x}(k)$ using an observer with gain $\mathbf{L}(k)$. Denoting by $\mathbf{P}$ the observer's error covariance, we have the following recursion:

$$\mathbf{P}(k+1) = (A - \mathbf{L}C)\mathbf{P}(k)(A - \mathbf{L}C)' + \mathbf{L}(\mathbf{b}R + R_0)\mathbf{L}' + \Sigma \ ,$$

where R, R_0 and Σ denote the covariance matrices of $\boldsymbol{\omega}$, $\mathbf{v}$ and $\mathbf{w}$, respectively. Initially, to keep our exposition simple, we assume that the covariance R is known and time-invariant.

We adopt once more the framework of average cost minimization for Markov decision processes. Define the one-step cost

$$c(P, b) = \operatorname{tr} QP + \lambda\phi(b) \ ,$$

where Q is some positive definite matrix, $\lambda > 0$ is a weight and $\phi(b)$ is the communication cost associated with the control b, which will depend on the method used to control capacity.

The corresponding ACOE with relative value function V and cost ϱ is given by

$$\varrho + V(P) = \min_{b,L} \{c(P, b) + V((A - LC)P(A - LC)' + L(bR + R_0)L' + \Sigma)\} \ .$$

Since solving for optimal controls for this type of system is computationally hard, we try to explore the linearity of the system to design our protocols. A possible direction is to find the value function of the form $V(P) = \operatorname{tr} HP$, with H positive definite, that minimizes a bound on the average cost. To this purpose, we can write the ACOI:

$$\varrho + \operatorname{tr} HP \geq \min_{b,L} \{c(P, b) + \operatorname{tr} H((A - LC)P(A - LC)' + L(bR + R_0)L' + \Sigma)\} \ . \tag{7.1}$$

The optimal L for this value function is given by the Kalman gain $L = APC'(CPC' + bR + R_0)^{-1}$. Replacing the optimal L, we obtain

$$\varrho + \operatorname{tr} HP \geq \min_{b} \left\{c(P, b) + \operatorname{tr} H(APA' - APC'(CPC' + bR + R_0)^{-1}CPA' + \Sigma)\right\} \ .$$

We want to select H to minimize the bound ϱ. We can rewrite this problem as the tri-level optimization problem:

$$\min_{H \geq 0} \max_{P \geq 0} \min_{b} \left\{c(P, b) + \operatorname{tr} H(APA' - P - APC'(CPC' + bR + R_0)^{-1}CPA' + \Sigma)\right\} \ .$$

Using the Schur complement, we can rewrite this problem as

$$\min_{H\geq 0}\max_{P\geq 0}\min_{b}\max_{K}\{c(P,b)+\operatorname{tr} HK)\}$$

subject to

$$\begin{bmatrix} APA'-P+\Sigma-K & APC' \\ CPA' & CPC'+bR+R_0 \end{bmatrix}\geq 0\ .$$

In general, multilevel optimization is a hard problem. We relax this problem by changing the order of optimization:

$$\min_{b}\min_{H\geq 0}\max_{P\geq 0}\max_{K}\{\operatorname{tr}(HK+QP)+\lambda\phi(b)\}$$

subject to

$$\begin{bmatrix} APA'-P+\Sigma-K & APC' \\ CPA' & CPC'+bR+R_0 \end{bmatrix}\geq 0\ . \tag{7.2}$$

If we fix b, we can regard the remaining optimization problem as the first step in a policy iteration algorithm. Hence, this optimization problem can be interpreted as finding the constant policy that minimizes the cost obtained in the next step of policy iteration.

Provided that $\phi(b)$ is convex, the advantage of the above problem is that it is jointly convex in b and H and jointly concave (linear) in P and K. Therefore, this is a minimax convex-concave problem that can be solved efficiently using convex programming techniques. However, we are mainly interested in the case the set of admissible actions b is discrete. To deal with this case, we use the following proposition.

Proposition 10. *Suppose that the pair (A,C) is detectable and that Σ and R_0 are positive definite. Let b^*, P^* and H^* denote the solutions to problem (7.2).*

Then, b^* and P^* are also solutions to the problem

$$\min_b \max_{P \geq 0} \{\operatorname{tr} QP + \lambda\phi(b)\}$$

subject to

$$\begin{bmatrix} APA' - P + \Sigma & APC' \\ CPA' & CPC' + bR + R_0 \end{bmatrix} \geq 0 \ ,$$

and H^* satisfies the Lyapunov equation

$$(A - L^*C)'H^*(A - L^*C) - H^* + Q = 0 \ , \tag{7.3}$$

where $L^* = AP^*C'(CP^*C' + b^*R + R_0)^{-1}$.

Proof. Since (A, C) is detectable and $\Sigma, R_0 > 0$, we have that the constraint in (7.2) for P is strictly feasible and the constraint set is bounded. Therefore, we can use the minimax theorem to rewrite (7.2) as

$$\min_b \max_{P \geq 0} \max_K \min_{H \geq 0} \{\operatorname{tr}(HK + QP) + \lambda\phi(b)\}$$

subject to

$$\begin{bmatrix} APA' - P + \Sigma - K & APC' \\ CPA' & CPC' + bR + R_0 \end{bmatrix} \geq 0 \ .$$

From this new form we can see that the optimal solution must have K positive semidefinite, otherwise $\inf_H \operatorname{tr} KH = -\infty$. When $K \geq 0$, $\inf_H \operatorname{tr} KH = 0$. Therefore, we can simplify this problem to

$$\min_b \max_{P \geq 0} \{\operatorname{tr} QP + \lambda\phi(b)\}$$

subject to

$$\begin{bmatrix} APA' - P + \Sigma & APC' \\ CPA' & CPC' + bR + R_0 \end{bmatrix} \geq 0 \ .$$

To calculate H^*, we use the Karush-Kuhn-Tucker conditions for the original cost

$$\operatorname{tr} H(APA' - P - APC'(CPC' + bR + R_0)^{-1}CPA' + \Sigma + KP) + \lambda\phi(b) \ .$$

Note first that P^* must be positive definite since $\Sigma > 0$. Therefore, the Karush-Kuhn-Tucker condition for the derivative with respect to P gives

$$\operatorname{tr} H(A\Delta_P A' - \Delta_P - A\Delta_P(L^*C)' + L^*C\Delta_P(L^*C)' - L^*C\Delta_P A' + K) = 0, \ \forall \Delta_P \ , \tag{7.4}$$

where Δ_P denotes first variations in P. We obtain (7.3) by rearranging the terms and using the fact that (7.4) must hold for all Δ_P. □

One should note that the use of the linear value function $\operatorname{tr} HP$ implies a threshold type policy when $c(P, b)$ is concave in b. Indeed, we have that the minimization in (7.1) is concave in b, which implies that the optimal action is either the minimum or the maximum b available. On one side, this provides a simple implementation rule. On the other side, not using the intermediary values of b may give worse suboptimal protocols.

Next, we want to consider a noise model where, at every time instant, $\boldsymbol{\omega}(k)$ has covariance R_i with probability p_i, $i = 1, 2$. The ACOE now becomes

$$\varrho + \operatorname{tr} HP \leq \min_{b, L_i} \Big\{ c(P, b) + \operatorname{tr} \sum_{i=1,2} p_i H((A - L_iC)P(A - L_iC)' + L_i(bR_i + R_0){L_i}' + \Sigma) \Big\} \ .$$

We can repeat the design procedure above to obtain a relative value function $V = \operatorname{tr} H^*P$, where H^* satisfies

$$\sum_{i=1,2} p_i(A - L_i^*C)'H^*(A - L_i^*C) - H^* + Q = 0 \ , \tag{7.5}$$

where $L_i^* = AP^*C'(CP^*C' + b^*R_i + R_0)^{-1}$, $i = 1, 2$ and b^*, P^* solve

$$\min_b \max_{P \geq 0} \{\operatorname{tr} QP + \lambda\phi(b)\}$$

subject to

$$\begin{bmatrix} APA' - P + \Sigma & \sqrt{p_1}APC' & \sqrt{p_2}APC' \\ \sqrt{p_1}CPA' & CPC' + bR_1 + R_0 & 0 \\ \sqrt{p_2}CPA' & 0 & CPC' + bR_2 + R_0 \end{bmatrix} \geq 0 \ .$$

7.2 Numerical Examples

We tested the proposed protocol design with a scalar example where $A = 2$, $C = 1$, $\Sigma = 3$, $Q = 1$, $p_1 = 0.85$, $p_2 = 0.15$, $R_1 = 0.1$, $R_2 = 3$, $R_0 = 0.1$. The control actions were on the set $\{10^{-2}, 10^{-1.4}, \ldots, 10^{3.4}, 10^{4}\}$ and the communication cost was $\phi(b) = -\log b + 4$.

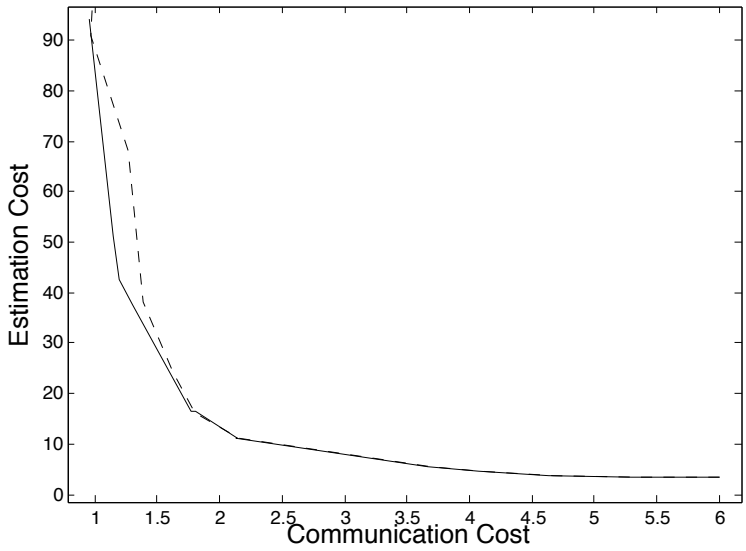

Figure 7.1: Pareto set for capacity scheduling.
Pareto frontier (solid); policy obtained from Proposition 10 (dashed).

The Pareto frontier in Figure 7.1 was obtained by varying λ from 0.001 to 1000. This figure shows that communication costs can be decreased significantly without much performance degradation in the estimation error. From this figure and Figure 7.2, we see that the policy given by Proposition 10 performs close to the optimal and better than the greedy policy (which corresponds to using a value function $\operatorname{tr} QP$). It is also important to notice that the upper bound ϱ predicted by Proposition 10 is quite conservative, being even worse than the greedy policy.

We consider next a two-dimensional example (for which the computation of optimal policies becomes harder) where $A = [1\ 1; 0\ 1]$, $C = [1\ 0]$, $\Sigma = [3\ 0; 0\ 3]$,

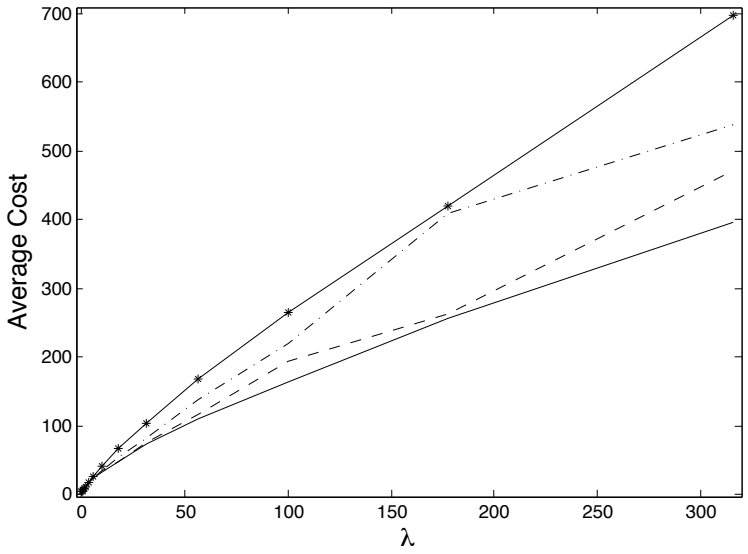

Figure 7.2: Cost obtained by capacity scheduling for different weights. Optimal policy (solid); policy obtained from Proposition 10 (dashed); cost predicted by Proposition 10 (*); greedy policy (dot-dashed).

$Q = [1\ 0; 0\ 1]$, $p_1 = 0.85$, $p_2 = 0.15$, $R_1 = 0.1$, $R_2 = 3$, $R_0 = 0.1$. The control actions were on the set $\{10^{-2}, 10^{-1.4}, \ldots, 10^{3.4}, 10^{4}\}$ and the communication cost was $\phi(b) = -\log b + 4$. As in the previous example, Figure 7.3 shows that a good tradeoff between communication and estimation cost can be achieved.

In this example, the policy given by Proposition 10 outperforms the greedy policy by a higher margin than in the past example, as shown by Figure 7.4 .

7.3 Comments and Open Problems

New directions on the material presented in this chapter include improving the proposed suboptimal policy or giving guarantees of how close to the optimal it performs. An interesting possibility that also explores the linearity of the plant is to consider piecewise linear value functions.

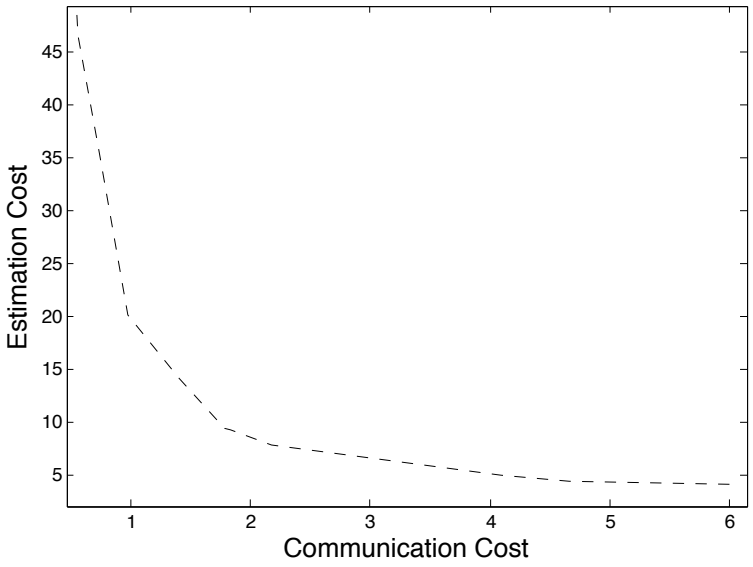

Figure 7.3: Pareto set for capacity scheduling in the 2-D example. Policy obtained from Proposition 10 (dashed).

As in our redundant transmission protocols, calculating the error covariance may be a computational burden to sensors. Hence, it would be valuable to develop simpler protocols. Unfortunately, the information to be considered is not as simple as counting the number of consecutive drops. A simple possibility would be given by protocols that depend on some time average of the noise power.

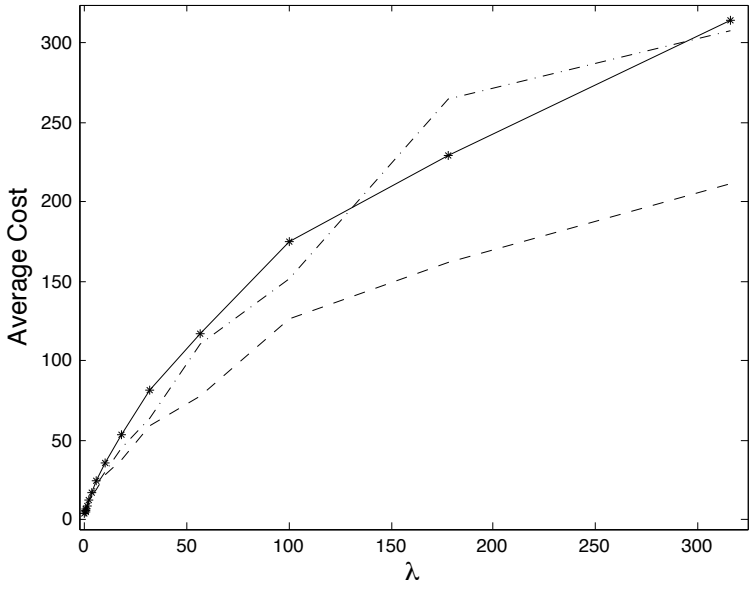

Figure 7.4: Cost obtained by capacity scheduling for different weights in the 2-D example.
Policy obtained from Proposition 10 (dashed); cost predicted by Proposition 10 (*); greedy policy (dot-dashed).

Bibliography

[1] T. Ahmed, T. S. Shimizu, and R. Stocker. Bacterial chemotaxis in linear and nonlinear steady microfluidic gradients. *Nano Letters*, 10(9):3379–3385, 2010.

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

[3] F. Aluffipentini, V. Parisi, and F. Zirilli. Global optimization and stochastic differential-equations. *Journal of Optimization Theory and Applications*, 47(1):1–16, 1985.

[4] A. Anta and P. Tabuada. Self-triggered stabilization of homogeneous control systems. In *American Control Conference, 2008*, pages 4129–4134. IEEE, 2008.

[5] J. Apostolopoulos. Reliable video communication over lossy packet networks using multiple state encoding and path diversity. In *Proceedings of SPIE, the International Society for Optical Engineering*, volume 4310, pages 392–409. Society of Photo-Optical Instrumentation Engineers, 2001.

[6] A. Arapostathis, V. Borkar, E. Fernández-Gaucherand, M. Ghosh, and S. Marcus. Discrete-time controlled Markov processes with average cost criterion: a survey. *SIAM Journal on Control and Optimization*, 31(2):282–344, 1993.

[7] K. Astrom and B. Bernhardsson. Comparison of Riemann and Lebesgue sampling for first order stochastic systems (I). In *IEEE Conference on Decision and Control*, volume 2, pages 2011–2016. IEEE; 1998, 2002.

[8] J. Aubin and H. Frankowska. *Set-Valued Analysis*, chapter Differential inclusions. Birkhauser, 2008.

[9] R. Bachmayer and N. Leonard. Vehicle networks for gradient descent in a sampled environment. In *Proceedings of the 41st IEEE Conference on Decision and Control*, pages 112–117, December 2002.

[10] D. Baronov and J. Baillieul. Autonomous vehicle control for ascending/descending along a potential field with two applications. In *American Control Conference, 2008*, pages 678–683. IEEE, 2008.

[11] M. Batalin, M. Rahimi, Y. Yu, D. Liu, A. Kansal, G. Sukhatme, W. Kaiser, M. Hansen, G. Pottie, M. Srivastava, et al. Call and response: experiments in sampling the environment. In *Proceedings of the 2nd international conference on Embedded networked sensor systems*, pages 25–38. ACM, 2004.

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

[13] E. Ben-Jacob, I. Cohen, and H. Levine. Cooperative self-organization of microorganisms. *Advances in Physics*, 49(4):395–554, 2000.

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

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

[16] S. Block, J. Segall, and H. Berg. Adaptation kinetics in bacterial chemotaxis. *Journal of bacteriology*, 154(1):312, 1983.

[17] M. 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.

[18] F. Bullo, J. Cortés, and S. Martinez. *Distributed Control of Robotic Networks*. Princeton University Press, 2009.

[19] E. Burian, D. Yoerger, A. Bradley, and H. Singh. Gradient search with autonomous underwater vehicles using scalar measurements. In *Proceedings of the 1996 Symposium on Autonomous Underwater Vehicle Technology, 1996. AUV '96.*, pages 86–98, June 1996.

[20] J. Chakareski and B. Girod. Rate-distortion optimized packet scheduling and routing for media streaming with path diversity. In *Data Compression Conference, 2003. Proceedings. DCC 2003*, pages 203–212. IEEE, 2003.

[21] P. Chou and Z. Miao. Rate-distortion optimized streaming of packetized media. *IEEE Transactions on Multimedia*, 8(2):390–404, 2006.

[22] R. Cogill, S. Lall, and J. Hespanha. A Constant Factor Approximation Algorithm for Event-Based Sampling. *American Control Conference, 2007. ACC'07*, pages 305–311, 2007.

[23] O. L. V. Costa and F. Dufour. Stability and ergodicity of piecewise deterministic markov processes. *SIAM J. Control Optim.*, 47(2):1053–1077, 2008.

[24] T. Cover and J. Thomas. *Elements of information theory*. John Wiley and sons, 2006.

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

[26] A. Dekkers and E. Aarts. Global optimization and simulated annealing. *Math. Program.*, 50(3):367–393, 1991.

[27] J. Deuschel and D. Stroock. *Large deviations*. Academic Press, 1989.

[28] A. Dhariwal, G. Sukhatme, and 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.

[29] S. Di Cairano, K. Johansson, A. Bemporad, and R. Murray. Discrete and Hybrid Stochastic State Estimation Algorithms for Networked Control Systems. *Lecture Notes in Computer Science*, 4981:144, 2008.

[30] D. Djonin and V. Krishnamurthy. MIMO Transmission Control in Fading Channels: A Constrained Markov Decision Process Formulation With Monotone Randomized Policies. *IEEE Transactions on Signal Processing*, 55(10):5069–5083, 2007.

[31] O. do Valle Costa, M. Fragoso, and R. Marques. *Discrete-time Markov jump linear systems*. Springer Verlag, 2005.

[32] M. Dorigo and L. M. Gambardella. Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation*, 1(1):53–66, April 1997.

[33] D. Down, S. Meyn, and R. Tweedie. Exponential and uniform ergodicity of Markov processes. *The Annals of Probability*, pages 1671–1691, 1995.

[34] R. Erban and H. Othmer. From individual to collective behavior in bacterial chemotaxis. *SIAM Journal on Applied Mathematics*, 65(2):361–391, 2004.

[35] E. Feinberg and A. Shwartz. *Handbook of Markov decision processes: methods and applications*. Kluwer Academic Publishers, 2002.

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

[37] N. Ghods and M. Krstic. Multi-agent deployment around a source in one dimension by extremum seeking. In *American Control Conference (ACC), 2010*, pages 4794–4799. IEEE, 2010.

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

[39] A. Goldsmith. *Wireless communications*. Cambridge Univ Pr, 2005.

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

[41] D. Grünbaum. Advection–diffusion equations for generalized tactic searching behaviors. *Journal of Mathematical Biology*, 38(2):169–194, 1999.

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

[43] O. Hernández-Lerma and J. Lasserre. Approximation schemes for infinite linear programs. *SIAM Journal on Optimization*, 8:973, 1998.

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

[45] 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, Dec. 1995.

[46] A. Howard, M. Mataric, and G. Sukhatme. Mobile sensor network deployment using potential fields: A distributed, scalable solution to the area coverage problem. *Distributed autonomous robotic systems*, 5:299–308, 2002.

[47] O. Imer and T. Basar. Optimal estimation with limited measurements. *International Journal of Systems, Control and Communications*, 2(1):5–29, 2010.

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

[49] A. Julius, M. Sakar, E. Steager, U. Cheang, M. Kim, V. Kumar, and G. Pappas. Harnessing bacterial power in microscale actuation. In *IEEE International Conference on Robotics and Automation, 2009. ICRA'09.*, pages 1004–1009. IEEE, 2009.

[50] A. Kacelnik, J. Krebs, and C. Bernstein. The ideal free distribution and predator-prey populations. *Trends in Ecology & Evolution*, 7(2):50–55, 1992.

[51] Y. Kalinin, L. Jiang, Y. Tu, and M. Wu. Logarithmic sensing in Escherichia coli bacterial chemotaxis. *Biophysical journal*, 96(6):2439–2448, 2009.

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

[53] A. Karmokar, D. Djonin, and V. Bhargava. Optimal and suboptimal packet scheduling over correlated time varying flat fading channels. *Wireless Communications, IEEE Transactions on*, 5(2):446–456, 2006.

[54] E. F. Keller and L. A. Segel. Model for chemotaxis. *Journal of Theoretical Biology*, 30(2):225, 1971.

[55] I. Kontoyiannis and S. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *Annals of Applied Probability*, pages 304–362, 2003.

[56] I. Kontoyiannis and S. Meyn. Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electron. J. Probab*, 10(3):61–123, 2005.

[57] A. Lasota and M. Mackey. *Chaos, Fractals, and Noise. Stochastic Aspects of Dynamics*, volume 97 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1994.

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

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

[60] G. Lipsa and N. Martins. Optimal state estimation in the presence of communication costs and packet drops. In *47th Annual Allerton Conference on Communication, Control, and Computing, 2009. Allerton 2009.*, pages 160–169. IEEE, 2010.

[61] 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*, 30(9):854–860, September 1985.

[62] A. Matveev, H. Teimoori, and A. Savkin. Navigation of a non-holonomic vehicle for gradient climbing and source seeking without gradient estimation. In *American Control Conference (ACC), 2010*, pages 219–223. IEEE, 2010.

[63] C. Mayhew, R. Sanfelice, and A. Teel. Robust source-seeking hybrid controllers for autonomous vehicles. In *Proceedings of the 2007 American Control Conference. ACC '07*, pages 1185–1190, July 2007.

[64] C. Mayhew, R. Sanfelice, and A. Teel. Robust hybrid source-seeking algorithms based on directional derivatives and their approximations. In *47th IEEE Conference on Decision and Control, 2008. CDC 2008.*, pages 1735–1740. IEEE, 2009.

[65] R. Mesibov, G. Ordal, and J. Adler. The range of attractant concentrations for bacterial chemotaxis and the threshold and size of response over this range. *The Journal of General Physiology*, 62(2):203, 1973.

[66] A. R. Mesquita, J. P. Hespanha, and K. J. Åström. Optimotaxis: A stochastic multi-agent optimization procedure with point measurements. In M. Egerstedt and B. Mishra, editors, *Hybrid Systems: Computation and Control*, number 4981 in LNCS, pages 358–371. Springer-Verlag, Berlin, Mar. 2008.

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

[68] S. Meyn and R. Tweedie. Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, pages 518–548, 1993.

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

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

[71] T. Morse, S. Lockery, and T. Ferrée. Robust spatial navigation in a robot inspired by chemotaxis in Caenorhabditis elegans. *Adaptive Behavior*, 6(3-4):393, 1998.

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

[73] H. Othmer and T. Hillen. The diffusion limit of transport equations II: Chemotaxis equations. *SIAM Journal on Applied Mathematics*, 62(4):1222–1250, 2002.

[74] S. Pang and J. Farrell. Chemical plume source localization. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 36(5):1068–1080, October 2006.

[75] M. Rahimi, R. Pon, W. Kaiser, G. Sukhatme, D. Estrin, and M. Srivastava. Adaptive sampling for environmental robotics. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04.*, volume 4, pages 3537–3544. IEEE, 2004.

[76] H. Risken. *The Fokker-Planck Equation - Methods of Solution and Applications.* Springer Series in Synergetics. Springer-Verlag, Berlin, 1984.

[77] G. Roberts and R. Tweedie. Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms. *Biometrika*, 83(1):95–110, 1996.

[78] E. Şahin and A. Winfield. Special issue on swarm robotics. *Swarm Intelligence*, 2(2):69–72, 2008.

[79] M. Sharpe. *General theory of Markov processes.* Academic Pr, 1988.

[80] A. Singh, A. Krause, C. Guestrin, and W. Kaiser. Efficient informative sensing using multiple robots. *Journal of Artificial Intelligence Research*, 34(1):707–755, 2009.

[81] A. Singh, R. Nowak, and P. Ramanathan. Active learning for adaptive mobile sensing networks. In *The Fifth International Conference on Information Processing in Sensor Networks, 2006. IPSN 2006.*, pages 60–68. IEEE, 2006.

[82] J. Sousa, K. Johansson, J. Silva, and A. Speranzon. A verified hierarchical control architecture for co-ordinated multi-vehicle operations. *International Journal of Adaptive Control and Signal Processing*, 21(2-3):159–188, September 2007.

[83] D. Stroock. *Probability theory: an analytic view*. Cambridge Univ Pr, 1993.

[84] 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*, 9110, 1991.

[85] M. Trincavelli, M. Reggente, S. Coradeschi, A. Loutfi, H. Ishida, and A. Lilienthal. Towards environmental monitoring with mobile robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008. IROS 2008.*, pages 2210–2215. IEEE, 2008.

[86] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[87] Y. Tu, T. Shimizu, and H. Berg. Modeling the chemotactic response of Escherichia coli to time-varying stimuli. *Proceedings of the National Academy of Sciences*, 105(39):14855, 2008.

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

[89] O. Vega-Amaya. The average cost optimality equation: a fixed point approach. *Bol. Soc. Mat. Mexicana*, 9(1):185–195, 2003.

[90] S. Vempala. Geometric Random Walks: A Survey. *MSRI volume on Combinatorial and Computational Geometry*, 2005.

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

[92] Y. Xu. *Communication scheduling methods for estimation over networks*. PhD thesis, University of California, Santa Barbara, CA, USA, March 2006. Available at `http://www.ece.ucsb.edu/~hespanha/published`.

[93] Y. Xu and J. Hespanha. Optimal Communication Logics for Networked Control Systems. In *In Proc. of the 43rd Conf. on Decision and Contr*. IEEE, 2004.

[94] H. Yang and M. Alouini. Markov chains and performance comparison of switched diversity systems. *IEEE Transactions on Communications*, 52(7):1113–1125, 2004.

[95] T. Yi, Y. Huang, M. Simon, and J. Doyle. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the National Academy of Sciences of the United States of America*, 97(9):4649, 2000.

[96] D. Zarzhitsky, D. Spears, and W. Spears. Swarms for chemical plume tracing. In *Swarm Intelligence Symposium, 2005. SIS 2005. Proceedings 2005 IEEE*, pages 249–256. IEEE, 2005.

[97] E. Zeidler. *Nonlinear functional analysis and its applications*. Springer Verlag, 1985.

[98] C. Zhang, D. Arnold, N. Ghods, A. Siranosian, and M. Krstic. Source seeking with non-holonomic unicycle without position measurement and with tuning of forward velocity. *Systems & Control Letters*, 56(3):245–252, March 2007.

[99] M. Zorzi and R. Rao. Capture and retransmission control in mobile radio. *IEEE Journal on Selected Areas in Communications*, 12(8):1289–1298, 1994.

Appendices

Appendix A

Proof of Theorem 19

Our proof is based on Theorem 9 and it consists of finding a Lyapunov function V that satisfies a Foster-Lyapunov condition uniformly on the set of policies. The numbers (1) and (2) of our theorem follow directly from Theorem 9. Number (3) follows from Theorem 10 under the same assumptions.

Since B is equipped with the discrete topology, Assumption 4 is trivially verified. Next, we show that Assumption 3 also holds. Define the function $s(e,v) := 1 - p^v$ and the measure $\nu(C) = \int_C f(y)dy$ for $C \subset \mathbb{R}^n$ measurable. Then, we have from (6.18) that

$$P(C|e,v) \geq \nu(C)s(e,v) \ , \tag{A.1}$$

which satisfies Assumption 3 *(iii)*. Because of our restriction of the set of admissible actions in (6.19), we have that $\int s(e,\pi(e))\nu(de) > 0$ for all $\pi \in \Pi_0$. This satisfies Assumption3 *(v)*.

Let B_L denote the open ball with radius L centered at the origin in $\mathbb{R}^n$, let 1_{B_L} be its indicator function and $\nu_0 := \nu(B_L)$. For some positive definite matrix $K \in M^{n\times n}$, let α be a constant such that $\nu_0 + p^M(1-\nu_0) < \alpha < 1$ and such that there exists $H \in M^{n\times n}$, the unique positive definite solution of

$$\alpha^{-1}p^M A'HA - H = -K \ . \tag{A.2}$$

The existence of such α is guaranteed by our assumption that $a^2p^M < 1$, since this implies that $\alpha^{-1/2}p^{M/2}A$ is Schur for α close enough to 1.

We define the Lyapunov function

$$V(e) = e'He + a_0 1_{B_L}(e) + b_0 \ , \tag{A.3}$$

where a_0 is a constant to be determined and b_0 is set to

$$b_0 = \frac{a_0(\alpha - \nu_0) - d_1}{1 - \alpha} \tag{A.4}$$

and the constant d_1 is defined as

$$d_1 := \frac{\alpha}{p^M} \max_{\|e\| \leq L} e'((1 - p^M)H - K)e + \operatorname{tr} H\Sigma \ . \tag{A.5}$$

We assume that $a_0 > 0$ is large enough so that $b_0 > 0$, which implies $V > b_0 > 0$. By our choice of V, there exists a positive constant δ such that $\sup_{B(e)} c(e, v) < \delta V(e)$, which verifies Assumption 3 *(i)*.

Assumption 3 *(ii)*, namely, $\int V d\nu < \infty$, is also satisfied since the Gaussian distribution has finite second moments.

The last assumption to verify is Assumption 3 *(iv)*, which in our formalism can be written as

$$\int V(y)P(dy|e, \pi(e)) \leq \alpha V(e) + s(e, \pi(e)) \int V(y)\nu(dy) \tag{A.6}$$

for all $\pi \in \Pi_0$. This condition can be understood as a Lyapunov-Foster condition that is satisfied uniformly on the set of policies. To verify that (A.6) indeed holds, we define

$$\begin{aligned}
\Delta V(e) :&= \int V(y)P(dy|e, \pi)) - s(e, \pi) \int V(y)\nu(dy) \\
&= \int p^\pi V(y) f(y - Ae)dy \leq p^\pi e'A'HAe + p^\pi \operatorname{tr} H\Sigma + p^\pi a_0\nu_0 + p^\pi b_0 \\
&= p^{\pi - M}\alpha e'(H - K)e + p^\pi \operatorname{tr} H\Sigma + p^\pi a_0\nu_0 + p^\pi b_0 \ ,
\end{aligned}$$

where the last equality comes from (A.2) and the dependence of π on e was omitted to simplify the notation. From this it follows that

$$\begin{aligned}
\Delta V - \alpha V \leq \frac{\alpha}{p^M} e' \left((p^\pi - p^M)H - K\right) e + \\
p^\pi \operatorname{tr} H\Sigma + p^\pi a_0\nu_0 - \alpha a_0 1_{B_L}(e) + (p^\pi - \alpha)b_0 \ .
\end{aligned} \tag{A.7}$$

In the region $\|e\| < L$, we can upper bound the right-hand side of (A.7) by using the fact that $\pi(e) \geq 0$:

$$\Delta V - \alpha V \leq \frac{\alpha}{p^M} e' \left((1 - p^M)H - K\right) e + \operatorname{tr} H\Sigma + a_0(\nu_0 - \alpha) + (1 - \alpha)b_0 \leq 0 \ ,$$

where the last inequality is obtained from (A.4) and (A.5).

In the region $\|e\| \geq L$, we have that $\pi(e) = M$, which implies that the right-hand side of (A.7) decreases strictly with $\|e\|$ in this region, except for $\|e\| = L$, where it jumps. Thus, it only remains to investigate if (A.6) is satisfied at $\|e\| = L$. For $\|e\| = L$, the following upper bound can be obtained from (A.7):

$$\Delta V - \alpha V \leq -\alpha p^{-M} \lambda_{\min}(K) L^2 + p^M \operatorname{tr} H\Sigma + a_0 p^M \nu_0 + (p^M - \alpha) b_0 \ .$$

Replacing (A.4) we obtain

$$\Delta V - \alpha V \leq -\alpha p^{-M} \lambda_{\min}(K) L^2 + p^M \operatorname{tr} H\Sigma + \frac{a_0 \alpha (\nu_0 + p^M(1-\nu_0) - \alpha) + (\alpha - p^M) d_1}{1-\alpha} \ .$$

Since α was chosen to satisfy $(\nu_0 + p^M(1-\nu_0) - \alpha) < 0$, one can make a_0 large enough so that $\Delta V \leq \alpha V$ for $\|e\| \leq L$. With this we have verified (A.6). □