UNIVERSITY OF CALIFORNIA
SANTA CRUZ

UNCERTAINTY-ANTICIPATING STOCHASTIC OPTIMAL FEEDBACK CONTROL OF AUTONOMOUS VEHICLE MODELS

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

DOCTOR OF PHILOSOPHY

in

APPLIED MATHEMATICS AND STATISTICS

by

Ross P. Anderson

March 2014

The Dissertation of Ross P. Anderson is approved:

Professor Dejan Milutinović, Chair

Professor Herbert Lee

Professor Gabriel Elkaim

Professor William Dunbar

Professor Randal Beard

Professor Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

Abstract

Uncertainty-Anticipating Stochastic Optimal Feedback Control of Autonomous Vehicle Models

by

Ross P. Anderson

Control of autonomous vehicle teams has emerged as a key topic in the control and robotics communities, owing to a growing range of applications that can benefit from the increased functionality provided by multiple vehicles. However, the mathematical analysis of the vehicle control problems is complicated by their nonholonomic and kinodynamic constraints, and, due to environmental uncertainties and information flow constraints, the vehicles operate with heightened uncertainty about the team's future motion. In this dissertation, we are motivated by autonomous vehicle control problems that highlight these uncertainties, with particular attention paid to the uncertainty in the future motion of a secondary agent. Focusing on the Dubins vehicle and unicycle model, this dissertation proposes a stochastic modeling and optimal feedback control approach that anticipates the uncertainty inherent to the systems. We first consider the application of a Dubins vehicle that should maintain a nominal distance from a target with an unknown future trajectory, such as a tagged animal or vehicle. Stochasticity is introduced in the problem by assuming that the target's motion can be modeled as a

Wiener process, and the possibility for the loss of target observations is modeled using stochastic transitions between discrete states. A Bellman equation based on a Markov chain that is consistent with the stochastic kinematics is used to compute an optimal control policy that is shown to perform well both in the case of a Brownian target and for natural, smooth target motion. We also characterize the resulting optimal feedback control laws in comparison to their deterministic counterparts for the case of a Dubins vehicle in a stochastically varying wind. Turning to the case of multiple vehicles, we develop a method using a Kalman smoothing algorithm for multiple vehicles to enhance an underlying analytic control law. As a result the vehicles achieve a formation optimally and in a manner that is robust to the uncertainty caused by a lack of communication among the vehicles. Finally, as a way to deal with a key implementation issue of these controllers on autonomous vehicle systems, we propose a self-triggering scheme for stochastic control systems, whereby the time points at which the control loop should be closed are computed from predictions of the process in a way that ensures stability.

To my ever faithful parents.

Acknowledgments

It goes without saying that reaching this milestone — much less reaching it in a productive and enjoyable manner while holding on to at least a portion of my sanity — would not have been possible without the help and support of many. My heartfelt thanks goes out to these people. Dr.-Ing. Dejan Milutinović, I mention you first, as would only be proper, given how helpful and supportive you have been, both in person and behind the scenes, of me and our work together. Owing to your guidance and resolve, I feel I am ready for the next step, and I'm truly grateful. Next I would also like to express my appreciation to Dr. Gabriel Elkaim, Dr. William Dunbar, Dr. Herbert Lee, and Dr. Randal Beard, members of my Ph.D. dissertation committee, for their time and interest in evaluating this work. Efstathios Bakolas, Panagiotis Tsiotras, Georgi Dinolov, Andrew C. Moore, and Dimos V. Dimarogonas, I thank you for productive collaborative work (and a particular thanks to Dimos for a rewarding visit). Beyond my immediate collaborators, I am thankful to the other AMS faculty for shaping such a welcoming department. I also thank the administrative support, and in particular the ever-helpful Tracie Tucker for hiding her frustration with my incessant questions, and I promise to first check the website in the future. To my labmate and friend Jeremy Coupe, who shared in both the good and the bad days, thanks and good luck. Coupled with all the work, of course, was some play, and I must thank my friends and housemates for helping me to get away, relax, and clear my mind, with additional thanks to Chris Phelps for entertaining my trivial questions along the way, and to Katelyn White for so frequently

taking charge. However, I owe the greatest debt of gratitude to my parents. Every step I've taken has been met with much love and selfless support, and I thank you for this and for instilling in me the drive that brought be this far. So, I dedicate this work to you as a token of my appreciation.

Ross Anderson
Santa Cruz, November 2013

A Note on Previously Published Material

The text of this dissertation includes reprints of the following material:

- Anderson, R.P, and Milutinović, D., "A Stochastic approach to Dubins vehicle tracking problems'.' *IEEE Transactions on Automatic Control*, (in press).
- Anderson, R. P., Bakolas, E., Milutinović, D., and Tsiotras, P., "Optimal Feedback Guidance of a Small Aerial Vehicle in a Stochastic Wind", *AIAA Journal of Guidance, Control, and Dynamics*, vol. 36, issue 4, pp. 975-985, 2013.
- Anderson, R. P., and Milutinović, D., "Stochastic Optimal Enhancement of Distributed Formation Control Using Kalman Smoothers", Accepted for publication in *Robotica*.
- Anderson, R.P, Milutinović, D., and Dimarogonas, D.V., "Self-Triggered p-moment Stability for Continuous Stochastic State-Feedback Controlled Systems", submitted to the *IEEE Transactions on Automatic Control*.

  Initial results pertaining to this paper are available in the publication

  - Anderson, R.P., Milutinović, D., and Dimarogonas, D.V., "Self-Triggered stabilization of continuous stochastic state-feedback controlled systems," in Proceedings of the *European Control Conference*, Zürich, Switzerland, 2013.

This work was supported by University of California Santa Cruz Chancellor's Fellowship, NASA UARC Aligned Research Program (NAS2-03144), the Graduate Research Fellowship Program of the National Science Foundation (Award No. DGE-0809125), and by the Nordic Research Opportunity, which was jointly funded by the National Science Foundation and the Swedish Research Council (Vetenskapsrådet). Their support is gratefully acknowledged.

With regards to the authors of these preprints, the co-author Dejan Milutinović (dejan@soe.ucsc.edu) listed in these publication directed and supervised the research which forms the basis for the dissertation. The co-author Dimos V. Dimarogonas (dimos@kth.se) co-supervised research during a research visit by the Ross Anderson to KTH Royal Institute of Technology (Stockholm, Sweden) during Summer 2012. The co-author Efstathios Bakolas (bakolas@austin.utexas.edu) and his Post-Doctoral Adviser Panagiotis Tsiotras (tsiotras@gatech.edu) that appear in the second paper above provided the feedback control strategy for the case of the deterministic wind in that paper, whereas the author of this dissertation developed the results relevant to the stochastic case.

Chapter 1

Introduction

1.1 Motivation

Mobile multi-robot systems, comprised of multiple autonomous vehicles, are playing an increasingly important role in scientific studies, public safety, and industry [190]. The use of multiple robots introduces the advantages of scalability, reduction of a problem's overall computational load, and the ability to operate in environments that are hostile to humans [40]. Meanwhile, interactions among members in the group allow for a complex or spatially-expansive task to be distributed among a potentially large number of robots with both individual and group objectives. Environmental monitoring and sampling of environmental phenomena or contaminants [67, 113, 136, 232], for example, have demonstrated the significance and potential impact of multi-robot

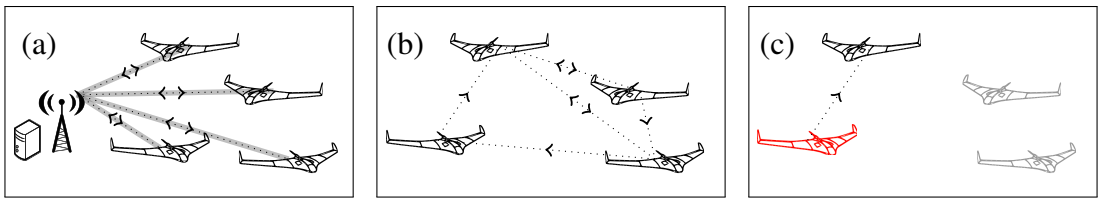


Figure 1.1: Illustration of architectures for autonomous vehicle control. (a) A centralized architectural model, where all communication (arrows) and commanded control (gray lines) are directed through a global supervisor. (b) A distributed design, where individual agents exchange relevant information along communication channels (arrows) but possess their own control algorithm. (c) Distributed control design without explicit communication, where each agent must possess a control algorithm capable of completing its objective using only information observable by that agent. In this case, the red vehicle observes one proximate vehicle (arrow) but is unaware of its future motion or may be unaware of the presence of other agents.

systems in aerial, aquatic, terrestrial, and subsoil arenas. Robotic infrastructures that exploit mobility have also proven useful in manufacturing, health care, defense, and other industrial areas [136].

An important engineering aspect of these systems revolves around the choice of architecture, which can take several forms, as shown in Fig. 1.1. On one side of the spectrum, in the *centralized* design paradigm, a global supervising algorithm monitors, plans, and controls the teams of vehicles during their task (Fig. 1.1a). However, in some circumstances, a centralized controller may not be feasible, or one may wish to take advantage of the distributed resources provided by each of the individual robots. Moving away from a centralized framework, one arrives at a distributed architectural model, in which autonomous vehicles act as interacting subsystems (Fig. 1.1b). This system architecture often relies on an underlying network of communication channels to propagate information among the vehicles.

The distributed architectural model has become increasingly popular in many real word applications, as it provides advantages from both a theoretical and practical standpoint. However, some criticism of the distributed design paradigm has focused on whether complex engineering systems with many degrees of freedom can be adequately controlled in such a manner due to, for instance, information structure constraints [277] that limit information propagation among the vehicles comprising the network, and due to the subsequent uncertainty in the operation of the team. For example, in a team of underwater autonomous vehicles, communication is often realistic only when the vehicles intermittently surface [6]. Additionally, information exchange among a team of terrestrial robots may be limited by transmission power and communication interference, for example.

This criticism may be more cogent in the case of a complete failure in the communication channels, or when the control strategy is designed *a priori* without communication (Fig. 1.1c), which can reduce system overhead and in some applications can be beneficial to system performance [31]. In this limiting case, the information available to one autonomous vehicle is explicitly restricted to that which it can immediately observe, and the notion of an underlying network among vehicles is less tangible. The vehicle could have knowledge of its current state, its local surroundings, and the observable state information of proximate robots (neighbors), for example. However, without further information exchange, the robot may be unaware of the objectives of neighboring robots, the neighboring robots' control algorithms, the inputs to these algorithms,

or even the presence of certain robots.

Consequently, from the perspective of one vehicle, the future motion of a neighboring robot is highly uncertain. This uncertainty is an inherent property of distributed mobile robotic systems, and is present alongside other sources of uncertainty in robotics identified in [259]. These include unpredictable and highly dynamic environments; noisy or range-limited sensors; failure and unpredictable behavior of robots and robotic components; inadequate realism of the abstraction provided by mathematical models; and issues caused by approximations made during the control design. Tantamount to the ability for an autonomous vehicle to achieve its goals is the capacity to *anticipate* and react to these sources of uncertainty.

In this dissertation, we are motivated by autonomous vehicle control problems that highlight these uncertainties, with particular attention paid to the uncertainty in the future motion of a secondary agent. In the case of just one vehicle, we first focus on the application of a small unmanned aerial vehicle (UAV) that should track from a nominal distance a target, such as a tagged animal or a vehicle under protection or surveillance. Since the future motion of the target is unknown, our model includes this uncertainty alongside the nonlinear motion of the tracking vehicle. This problem is further motivated by the possibility for stochastic loss of target observation due to sensor obstruction, for example. In this case, the control strategy for the tracking vehicle upon observation loss is necessarily based on outdated information, and we include this in

our analysis. A related application[1] is the navigation of a small UAV toward a target or waypoint in an unpredictable, turbulent wind. Motivated by the potential for a significant effect of wind on the vehicle motion, we are interested in addressing how the control strategy for the vehicle should account for the uncertainty in the wind.

In the case of multiple vehicles, we extend our analysis to the application of formation control, which is thought to increase a robotic team's efficiency and performance. In this problem, each vehicle should nominally distance itself from proximate vehicles using its individual observations. Formation control can therefore essentially be interpreted as a coupling of the previously-mentioned tracking problems but, with the added degrees of freedom describing multiple vehicles, can be much more challenging of a control problem.

The autonomous vehicles used in the aforementioned applications are typically nonholonomic[2] [48, 71, 155]. For example, an automobile is constrained by the direction of its wheels, and a UAV must maintain a forward speed and cannot turn too sharply. In this dissertation, we focus on control design for unicycle models and Dubins vehicle models [82]. The former is commonly studied as an abstraction for wheeled vehicles, while the Dubins vehicle model provides a good approximation to the motion of a fixed-speed UAV with a bounded turning rate [204–206] or a cruising automobile, for example. Although these models are rather idealized for practical vehicles, they are

[1] Based on our modeling choices, it will turn out a specific case of this problem is kinematically equivalent to the previously mentioned tracking problem, and could further describe the motion of an underwater autonomous vehicle in uncertain ocean currents [9]

[2] In other words, they have (non-integrable) constraints on certain types of motion.

nonlinear and take into account kinematic limitations of the vehicles. As such, these models simultaneously complicate the control problem solution and give rise to a rich set of behaviors, especially when acting in response to the uncertain motion of a proximate autonomous vehicle or target.

Control problems involving these models have enjoyed considerable treatment in the past, although they are usually approached with either Lyapunov function-based stability analyses or through optimal control or optimization of deterministic models. The main approach used in this dissertation, however, is stochastic optimal feedback control[3] based on the Hamilton-Jacobi-Bellman (HJB) partial differential equation (PDE) [36]. This approach lends several advantages to the problems at hand, but also raises questions on control computation and implementation that demand treatment in this dissertation. Before addressing the motivation for this approach and its consequences in detail, however, we first turn to the scope of this dissertation and its goals.

1.2 Thesis Contributions

The research presented in this dissertation pursues four main objectives. The primary problem studied revolves around a modeling and control approach for a nonlinear agent (specifically, a Dubins vehicle or unicycle) to achieve distance-based goals with respect

[3]To clarify, the "stochastic" in stochastic optimal feedback control refers to the model being controlled and not the (deterministic) mapping from state to control. All control laws presented in this dissertation are deterministic mappings.

to other agents or targets in the presence of stochasticity. We aim to apply existing, and develop new, computational methods for stochastic optimal feedback control of nonlinear stochastic models describing autonomous vehicles. In doing so, we focus on and control for the uncertainty due to the target's or neighboring vehicle's unknown future motion, the effects of outdated state information, and the effects of uncertain environmental phenomena (e.g., wind) on vehicle motion.

Many previous efforts along these fronts develop stabilizing feedback control laws for deterministic autonomous vehicle systems that, due to their analytic forms, may be considered more intuitive, or transparent, than the numerical controllers computed herein. To this end, the second goal of this dissertation seeks to describe how the numerical control laws that we compute relate to their analytic analogues. For the case of Dubins vehicle motion in a stochastic wind, for example, the minimum-time optimal feedback control in the limit of no noise, i.e., without stochastic wind, has a known analytical form, and we aim to characterize how it should change to remain optimal in the presence of the stochasticity. In addition, we will analyze the extent to which the controller performance improves when accounting for stochasticity. In the limit of light wind or turbulence, for instance, the performance increase when accounting stochasticity may not be significant and could be outweighed by the benefits of an analytical control law. Is it really necessary, then, to compute the stochastic optimal feedback control? This will be examined as part of the second objective. However, we should mention at this point that the efforts in this direction are largely application-

dependent, and conclusions are not intended to have scope beyond the problems at hand.

For some autonomous vehicle problems, analytical, deterministic optimal controllers have not previously been developed, often due, in part, to a large dimensional state space. In the formation control problem, in particular, although deterministic optimal controllers are not available, many non-optimal, stabilizing feedback controls that (eventually) bring about a vehicle formation have been developed. The resulting vehicle behaviors are qualitatively appealing, and they can be shown to guarantee important but complex system requirements like collision avoidance, for example. However, the vehicle trajectories are non-optimal (to a known cost function[4]) and are not designed with stochasticity in mind. One can then ask — similarly to the stochastic wind case — what additional feedback control input is necessary to drive the vehicles into a formation *optimally* and in a manner that is *robust to the uncertainty* induced by the distributed nature of the system. This gives rise to the notion of a stochastic optimal enhancement to stabilizing controllers for large-dimensional autonomous vehicle systems, and forms the basis of our third objective. However, the many degrees of freedom in these systems render traditional methods of computing an optimal feedback control, i.e., a numerical solution to the HJB equation, impossible. We therefore turn to the so-called path integral (PI) approach to stochastic optimal control and show how it is appropriate for the formation control problem with unicycle vehicle models. We further develop

[4]A stabilizing control law could, in some instances, be optimal with respect to some cost function [180], but is usually not designed with this in mind.

a new connection between the path integral control approach and Kalman smoothing algorithms that allows the optimal (enhancing) feedback control to be computed in real-time by each autonomous vehicle. While the number of degrees of freedom prohibits any general remarks to be made about how the underlying stabilizing controllers compare to the stochastic optimal feedback control, we advocate for the techniques presented herein as a new tool for the analysis and improvement of stabilizing controllers for distributed autonomous vehicle systems.

This dissertation next identifies and subsequently addresses a key challenge that arises during the implementation of control laws for stochastic autonomous vehicle systems. A autonomous vehicle control law, be it analytic or computationally-generated, it is often implemented on a micro-controller or other digital controller on board the autonomous vehicle. However, this practice requires a choice of the time points at which the controller should be updated. A modern trend in real-time control system design is self-triggered control, whereby these times are chosen using predictions based on the most recent observation in a way that can extend the length of time between updates while retaining the stability of the closed-loop system. Self-triggered control design for stochastic feedback control systems has only recently begun to receive attention in literature. Perhaps surprisingly, as this is not reported elsewhere, certain methods of computing a numerical stochastic optimal feedback control already offer a subtle notion of self-triggered update times required to keep stability, as discussed in Chapter 3. However, we are not aware of similar results for the analytic case. The fourth objective,

therefore, is to develop a self-triggered control scheme for (analytic) stochastic control systems. Particular attention is paid to the example of an autonomous wheeled cart in the presence of environmental uncertainties.

A good portion of this dissertation is motivated by and pertains to the "analytic *versus* numerical," "stabilizing *versus* optimal," and "deterministic *versus* stochastic" feedback control debates, but the issues raised here are by no means exhaustive. Admittedly, many applications in autonomous vehicle control systems would be more appropriately tackled by other means. It is envisioned, however, that this dissertation will provide evidence for some of the benefits to the numerical, optimal, and stochastic feedback control approach as applied to nonlinear, autonomous vehicle systems in the presence of uncertainties. Moreover, it is intended to help elucidate the role of noise in designing feedback control policies for autonomous vehicle teams.

1.3 Outline

This dissertation is organized as follows. Chapter 2 summarizes the state of available literature. Chapter 3 introduces some of the preliminary modeling and mathematical concepts and tools used in this dissertation. In doing so, we also present a road map of the remainder of the dissertation and further motivate the inclusion of the results in the sequel. Chapters 4-7 consist of appended papers detailing research results. Brief summaries of each of these papers appear below. Chapter 8 concludes the dissertation with a summary of remarks and directions for future research.

- **Chapter 4: A Stochastic Approach to Dubins Vehicle Tracking Problems**

  An optimal feedback control is developed for fixed-speed, fixed-altitude Unmanned Aerial Vehicle (UAV) to maintain a nominal distance from a ground target in a way that anticipates its unknown future trajectory. Stochasticity is introduced in the problem by assuming that the target motion can be modeled as Brownian motion, which accounts for possible realizations of the unknown target kinematics. Moreover, the possibility for the interruption of observations is included by assuming that the duration of observation times of the target is exponentially distributed, giving rise to two discrete states of operation. A Bellman equation based on an approximating Markov chain that is consistent with the stochastic kinematics is used to compute an optimal control policy that minimizes the expected value of a cost function based on a nominal UAV-target distance. Results indicate how the uncertainty in the target motion, the tracker capabilities, and the time since the last observation can affect the control law, and simulations illustrate that the control can further be applied to other continuous, smooth trajectories with no need for further computation.

- **Chapter 5: Optimal Feedback Guidance of a Small Aerial Vehicle in the Presence of Stochastic Wind**

  The navigation of a small unmanned aerial vehicle is challenging due to a large influence of wind to its kinematics. When the kinematic model is reduced to two dimensions, it has the form of the Dubins kinematic vehicle model. Con-

sequently, this paper addresses the problem of minimizing the expected time required to drive a Dubins vehicle to a prescribed target set in the presence of a stochastically varying wind. First, two analytically-derived control laws are presented. One control law does not consider the presence of the wind, whereas the other assumes that the wind is constant and known a priori. In the latter case it is assumed that the prevailing wind is equal to its mean value; no information about the variations of the wind speed and direction is available. Next, by employing numerical techniques from stochastic optimal control, feedback control strategies are computed. These anticipate the stochastic variation of the wind and drive the vehicle to its target set while minimizing the expected time of arrival. The analysis and numerical simulations show that the analytically-derived deterministic optimal control for this problem captures, in many cases, the salient features of the optimal feedback control for the stochastic wind model, providing support for the use of the former in the presence of light winds. On the other hand, the controllers anticipating the stochastic wind variation lead to more robust and more predictable trajectories than the ones obtained using feedback controllers for deterministic wind models.

- **Chapter 6: Stochastic Optimal Enhancement of Distributed Formation Control Using Kalman Smoothers**

  Beginning with a deterministic distributed feedback control for nonholonomic vehicle formations, we develop a stochastic optimal control approach for agents

to enhance their non-optimal controls with additive correction terms based on the Hamilton-Jacobi-Bellman equation, making them optimal and robust to uncertainties. In order to avoid discretization of the high-dimensional cost-to-go function, we exploit the stochasticity of the distributed nature of the problem to develop an equivalent Kalman smoothing problem in a continuous state space using a path integral representation. Our approach is illustrated by numerical examples in which agents achieve a formation with their neighbors using only local observations.

- **Chapter 7: Self-Triggered p-moment Stability for Continuous Stochastic State-Feedback Controlled Systems**

  Event-triggered and self-triggered control, whereby the times for controller updates are computed from either current or outdated sampled data, have recently been shown to reduce the computational load or increase task periods for real-time embedded control systems. In this work, we propose a self-triggered scheme for nonlinear controlled stochastic differential equations with additive noise terms. We find that the family of trajectories generated by these processes demands a departure from the standard deterministic approach to event- and self-triggering, and, for that reason, we use the statistics of the sampled-data system to derive a self-triggering update condition that guarantees p-moment stability. We show that the length of the times between controller updates as computed from the proposed scheme is strictly positive and provide related examples.

Chapter 2

Related Work

In this chapter, we identify and describe some previous research efforts on topics that are relevant to this dissertation. We begin with a review of some important results in the field of control for autonomous robot path and motion planning, and in particular, we focus on results related to Dubins vehicle models and unicycle models. This includes a review of some results for single vehicles with individual objectives; single vehicles in the presence of wind or ocean currents, and the minimum-time control problem; and multi-vehicle problems, including formation control. With the formation control problem in mind, we subsequently review strategies for optimal control for systems with large dimensional state spaces with a focus on the path integral approach. Finally, we review some results for self-triggered stabilization, with particular attention paid to stochastic control systems. Considering the breadth of the topics touched upon by this dissertation, the purpose of this section is to sketch the current state of the field with the intention of revealing the need for the new results presented in this dissertation.

2.1 Autonomous Vehicle and Dubins Vehicle Control

There are a great deal of results in the literature for autonomous vehicle control. This is likely caused more by the variety of vehicles and the particular applications involved than the advances in nonlinear control theory. By far, one of the most common control design technique involves the use of Lyapunov functions — defined in the next chapter — to guarantee stability of the closed-loop system toward some goal. Some prototypical examples along these lines are [216], which controls a wheeled cart, the paper by Aicardi *et al.* [5], which steers a unicycle-like vehicle[1] to a destination, and the references [155, 231], each of which guides a car-like vehicle to park. Another common goal is for the vehicle to track a known reference trajectory, either by converging upon and then matching its position [34, 35] or by maintaining a standoff distance [65].

One of the more popular vehicle models in nonholonomic path planning and control is that which moves in the direction of its heading, with a fixed speed, and is able to change the direction of its velocity vector with a bounded turning rate. This gives rise to a good first approximation to fixed-speed, fixed-altitude UAVs [39] and to cruising autonomous ground or marine vehicles that cannot turn too quickly. Finding the shortest paths taken by this kinematic model is intrinsically related to the problem of finding the shortest planar paths of bounded curvature, which was addressed by L. E. Dubins in 1957 [82] and is usually referred to as the Dubins problem (or the Markov-Dubins

[1] A vehicle that moves in the direction of its heading angle with controlled turning rate and with either a controlled forward speed or controlled acceleration, depending on the application.

problem due to the initial contributions of the Russian mathematician [243]). While Dubins did not actually speak of a kinematic vehicle, for simplicity and to keep with the terminology of existing works, we will also refer to the previously-mentioned kinematic model as the Dubins vehicle (DV) in this dissertation.

By examining the Dubins vehicle shortest path problem in the framework of time-optimal control and geometric control, Sussmann and Tang [242] and Boissonnat *et al.* [50] made rigorous the results of Dubins. Not only was this helpful for the problem of guiding a DV to a target in minimum time, but the presented techniques also made other optimal control problems involving the Dubins vehicle more approachable. These include optimal DV route tracking [237], optimal control to a straight line path [114], and other variants [236]. Further generalizations to the Dubins vehicle problem include that of Reeds and Shepp, which describes a vehicle with a variable forward speed [203, 242], extensions to three dimensions [64, 83, 159], and problems involving heterogeneous terrain [80, 217].

Owing to the similarity of the Dubins vehicle paths and UAV paths, the Dubins vehicle model has been embraced by the UAV community for path planning, guidance, and navigation, both for single UAV flight problems (see [39, 88, 119, 224] and the references therein, for example) and for multiple UAVs, which will be discussed later. However, the UAV guidance problem often presents some additional challenges not found in the Dubins problem. For problems with more complex kinematics, constraints, or objectives than the minimum-time Dubins problem, dynamic programming [200, 201]

and open-loop optimal control [78] have been proposed as methods to compute the UAV turning rate. In particular, the papers [78, 200, 201] compute optimal turning rate controls for a UAV team with respect to a ground target with a deterministic trajectory that is known *a priori*. Sensors on board the UAV [39] must also be taken into consideration, and while there are several works that aim to maintain focus of a target (see [79], for example), there are few results on control strategies when the target has been lost. Relevant to the problem of observation loss is the optimal search problem, but this is usually tackled in an open-loop fashion [182, 241].

One appealing characteristic of the aforementioned results dealing with the UAV / DV guidance problem and its variants is the prevalence of geometric segments that can often be used to piece together an optimal path. For example, one can characterize minimum-time planar Dubins vehicle paths in terms of combinations of straight line segments and curved segments [56]. In the presence of stochasticity, however, this may no longer be the case, and the design of control laws can be much more challenging. Nonetheless, there are a few ventures into the stochastic realm as well. Stochastic versions of Dubins vehicle control problems usually analyze DV routing through stochastically-generated targets [88, 89, 119, 222–224], since, conditioning on the position of the newly-generated target, the time-optimal path is known.

While path planning and motion planning for autonomous nonholonomic vehicles under uncertainty have received a fair deal of attention, results have largely concentrated on uncertainty in parameter values or on the challenges presented by an un-

structured environment [53, 98, 141, 246]. Kinematic uncertainties and environmental uncertainties that are more aptly described by stochastic noise [22] are less commonly incorporated into the control problem solution. Indeed, the role of stochastic noise in designing feedback policies for autonomous vehicles is still not well understood. Only a handful of relatively simple stochastic optimal control problems in the spirit of vehicle motion have known analytical solution [20, 43]. As such, stochastic vehicle control problems often rely on numerical methods [112, 201, 228, 229] and stability analyses [103, 151].

2.2 Minimum-Time Aerial Vehicle Control in an Uncertain Wind

A commonly-studied application related to the set of control problems discussed above is the problem of navigating an autonomous vehicle to a target or waypoint in minimum time[2] in the presence of a velocity field, such as wind or ocean currents. The effect of wind on the dynamics of the vehicle can be significant, as can the corresponding change in performance [156], especially in the case of a small aerial vehicle. This has led to a good number of works that attempt to compute or characterize minimum-time paths under the influence of local velocity fields, a problem frequently referred to as the Zermelo problem after [294].

It should be mentioned that there is some overlap of this problem with some of

[2]The nonlinear minimum-time problem has also been studied in generality in [23] and additionally with other vehicles in mind, such as autonomous underwater vehicles [208], differential drive robots [32], and omni-directional vehicles [33].

the studies mentioned in the previous section. In particular, the minimum time path problem in a wind field can be considered a generalization of the minimum-time path path problems for a Dubins vehicle (and its variants) in the absence of wind. In fact, some of the (sub-optimal) strategies presented for navigating in a partially-known or unknown wind field are identical to those of navigating without wind [27]. Moreover, the minimum-time path problem for a Dubins vehicle in an anisotropic environment also bares some semblance to the wind problem [80].

Stabilizing controllers for navigation in a constant wind field have been presented in [132, 215] and tested experimentally in [133, 179]. Similar approaches can be found in [55, 296], which use on-line estimates of the velocity of the wind field. The path planning problem in a constant wind field was framed as a numerical optimization problem in [60, 70, 193, 213, 238].

The minimum-time Dubins vehicle problem (that is, the problem of finding the minimum-time path) in the case of a constant wind was first posed in [165], and the control problem was solved numerically in [26, 254]. Also for the case of a constant wind, a full characterization of the optimal feedback control for the Dubins vehicle was presented in [24, 29]. Conditions for the existence and uniqueness of minimum-time trajectories in a deterministic wind field are described in [120].

A numerical algorithm that computes the minimum-time paths of the Dubins vehicle in the presence of a deterministic time-varying, yet spatially invariant, wind is presented in [167], and with temporally-constant but regionally-varying winds in [25].

The paper by Rhoads, Mezić, and Poje [208] considers a feedback control approach based on the HJB equation for minimum-time control of AUVs in a deterministic, time-varying flow field. The latter is somewhat similar to the work presented herein and identifies some of the key issues with the computational tractability of the HJB approach.

Common to the aforementioned references is the presence of a known, deterministic wind; little has been said about the case of minimum-time navigation in a stochastic wind. In [156], the statistics of wind from a weather forecast were incorporated into an approximate dynamic programming method to generate flight paths for long-distance travel. Strategies for Dubins vehicle navigation that are blind or partially blind to the wind and its statistics are presented in [27]. It remains to be seen in Chapter 5 how the minimum-time Dubins vehicle paths in the presence of a stochastically-varying wind field change from their deterministic counterparts.

2.3 Distributed Multi-Vehicle Coordinated Motion and Formation Control

While the study of multi-vehicle systems naturally extends that of single vehicles, it also brings about additional challenges in control. Cooperative control for autonomous vehicles (see [58, 59, 106, 190] for recent surveys) has been a significant source for research in the mobile robots community, particularly for spatially expansive tasks like coverage control [57, 160], for multi-vehicle consensus algorithms [206], and for theoretical and practical issues raised about the underlying vehicle communication network [136].

One of the first approaches to designing coordinated team motion was pioneered by C. Reynolds [207] and led to the behavior-based (or rule-based) design strategy [69, 143, 147, 276]. Agents are endowed with a set of behaviors, e.g., steer toward the direction of others' velocity vectors, and these rules often give rise to qualitatively attractive group trajectories. In some instances, analyses have been conducted on the resulting trajectories of the group to show convergence to a common state or goal (consensus) [47, 121], leading to, for example, flocking behaviors [184]. As an alternative to the direct design of the agent behaviors, many authors have taken a two-tiered approach in which an artificial potential function is first constructed, and then a control law that determines how a vehicle reacts to this potential is designed [94, 149, 150, 252] such that the vehicles arrive at one of the field's local minima. To avoid multiple local minima, navigation functions [209, 212] have also been used in literature for multi-vehicle consensus problems [76, 153, 250, 251]. However, unless care is taken in the design of these functions, oscillatory vehicle trajectories might arise. Game-theoretical approaches [54, 158] to distributed multi-agent control problems have also shown some success but impose a greater structure on the assumed objective or motion of a neighboring vehicle than what we seek to achieve in this work.

Problems where a particular spatial structure or pattern is desired of the vehicles are typically refered to as formation control, and these problems include formation acquisition, maintenance, and, in some instances, tracking of a reference trajectory by the group [107]. Formations of vehicles provide have attracted much attention in the UAV

community [38, 214], but also have shown the potential for application in commercial aviation [290], autonomous highway fleets [95], and for satellite formations [7]. Moreover, when each vehicle is equipped with a sensor or antenna, it can often be the case that a more powerful sensor or antenna can be synthesized when vehicles are in formation [8, 268].

Common strategies for multi-vehicle formation attainment and maintenance include the leader-follower strategy and the virtual structure or virtual leader paradigm. In the leader-follower strategy [72, 176, 177, 249], one vehicle acts as a leader, defining a clear reference trajectory, and other vehicles are tasked with maintaining a specified structure with respect to this leader. While this approach can significantly simplify control design and analysis, the unidirectional dependence on a leader can increase the effects of disturbances on the resulting performance [291]. Moreover, in some applications, it may be undesirable for all vehicles to rely on a single leader that may be prone to fault. The concept of a virtual leader removes the responsibility for a real vehicle leader and instead tasks a fictitious, “virtual" leader to provide a reference trajectory for other vehicles to track [86, 173]. However, for the team of vehicles to agree upon the state variables describing the leader (e.g., its position and heading), the level of communication among the vehicles may significantly increase. Similar to the virtual leader approach is the virtual structure approach, whereby an entire virtual formation is synthesized, and vehicles are tasked with tracking or attaining their respective positions in the virtual structure [109, 273, 293]. Artificial potential techniques are often

used in unison with these methods in order to provide collision avoidance (see, for example, [87, 145]).

UAV formation flight control [8], as with the previously-discussed applications, is commonly designed using notions of stability (see, for example, [38, 110, 132, 138]). However, unlike the single-vehicle case, the interaction topology of the vehicles plays a key role in the performance of the team, and this is usually tackled by incorporating results from graph theory into the stabilization problem [8]. Conditions under which the attainment of a stable formation is possible are developed in [73, 147, 148], and the references [95, 226] examine stability under the limiting case of no information transmission between vehicles.

Recent focus on optimizing aerial vehicle motion, either for fuel efficiency [221, 240] or for formation flight, has focused on LQR methods [227] and on receding-horizon control [52, 62, 84, 199, 284], the latter of which has proven successful at handling complex constraints like collision avoidance. In [230], the authors piece together optimal Dubins vehicle primitive segments (i.e., curves and straight lines) to plan the motion for the group. Dynamic programming has been used by S. Quintero *et al.* [200, 201] to optimally control the motion of multiple UAVs with respect to a ground target with known trajectory. A similar problem was posed and solved by X. Ding, A. Rahmani, and M. Egerstedt in [78] using open-loop control. Stochasticity is usually not included in these previous works. An exception is [201] in which the authors apply a dynamic programming approach using empirically-estimated UAV

transition probabilities to UAV flocking, but where the target (in this case, the leader UAV) has known control.

2.4 Stochastic Optimal Feedback Control with Many Degrees of Freedom and the Path Integral Approach

The most common approach to control for teams of autonomous vehicles seen in literature revolves around the use of Lyapunov functions and stability theory. This offers some advantages, but, as will be discussed in the following chapter (along with relevant references not included here), it can lessen the amount of influence a control designer has on the performance of the team. Our approach in this dissertation is stochastic optimal feedback control, which incorporates a cost function to be minimized. However, this method requires the solution of the HJB equation, and this usually entails a discretization of the state space. For systems with even just a few autonomous vehicles, the number of state variables of the system is already too large to be properly handled with the HJB equation on a discrete state space. This reflects the so-called "curse of dimensionality" [41] of dynamic programming.

Various methods have been proposed to overcome this issue. On the modeling side of the problem, one initial step is the reduction of the number of degrees of freedom of a model [81, 93] or the substitution of a model by a simpler abstraction that sufficiently captures its dynamics [189, 247]. In terms of the solution of the HJB equation, progress toward the solution of higher order problems has been made through

numerical methods that allow for parallelization [112], approximations to the solution to the HJB equation [175, 234, 261, 267], combinations of closed-loop with open-loop numerical techniques [68], and simplifications that occur for particular Hamiltonian structures [164].

A celebrated alternative to the backward recursion of the HJB equation and dynamic programming is approximate, or forward, dynamic programming (ADP) [45, 46, 195, 196]. (In communities outside control, especially artificial intelligence, the term reinforcement learning (RL) [245] is considered more fitting.) Without backward recursion of the cost, this method relies on the ability for forward-time sample trajectories to explore the state space and sufficiently capture the value of the possible paths in order to compute an optimal policy. Direct application of RL techniques for multi-agent systems [186] can be difficult, however, due to the multiple simultaneous learning processes, and convergence to an optimal policy may not occur [117]. Various algorithms to efficiently explore the state space using sampled trajectories have also been presented in, for example, [128, 134, 142, 210, 289].

A recent take on stochastic optimal control has examined the relation between the solutions to optimal control PDEs and the probability distribution of stochastic differential equations [292] in both the open-loop optimal control setting [169–171, 185] and in the closed-loop case. For the closed-loop case, through a clever logarithmic transformation [63] of the cost-to-go, W. H. Fleming discussed in a series of publications [96, 97] a class of nonlinear Hamilton-Jacobi-Bellman equations that can be linearized without

any loss of generality. From the Feynman-Kac equations, there then exists a representation of the solution to the linearized HJB in terms of the expected value of a function of samples of a forward-time, uncontrolled diffusion process. This PDE–diffusion duality allows one to solve stochastic optimal feedback control problems using the realizations of stochastic processes, a promising approach for systems with large-dimensional state space. In 2005, a connection between this representation and path integrals over the trajectories of the associated diffusion process was developed, giving rise to the field of path integral control [124]. The PI approach transforms the (continuous-time) HJB equation, which must be solved on a discrete state space, into an estimation problem on the distribution of (discrete-time) optimal trajectories in continuous state space.

This relation between optimal control and estimation is perhaps more familiar in the linear quadratic Gaussian (LQG) case [22], but this concept was also extended to the nonlinear quadratic Gaussian (NLQG) case [37] and to the fully nonlinear case in [263]. Along these lines, optimal estimation algorithms, including Kalman filter-based algorithms, have also been used in control problems directly [2, 269], but not as a method to compute an optimal feedback control using the PI approach. Instead, the computation of the optimal control in previous PI works has involved use of a Laplace approximation or various Markov chain Monte Carlo (MCMC) techniques [125]. Additionally, multi-agent systems have previously been studied using the PI approach [271, 272, 285, 286], but in these works, the agents cooperatively compute their control from a marginalization of the joint probability distribution of the group's trajectory, requiring explicit

communication among the agents. The relation of path integral control to concepts in statistical physics, reinforcement learning, and Markov decision processes has also received considerable mathematical treatment [85, 127, 256, 257].

2.5 Self-Triggering for Stochastic Control Systems

The communication constraints imposed on many autonomous vehicle systems often force a vehicle to use a control that is based on outdated information. Moreover, some on-line methods to compute a control (e.g., model predictive control [MPC]) can be slower than the time-scale of the vehicle dynamics or their control loops. Consequently, even if a vehicle could obtain data without delay, an update to its control algorithm could be based on outdated samples. The same issue can occur when a continuous-time control is reformulated as a sampled-data control in the digital domain. At a minimum, one should ensure that these outdated samples do not cause destabilization of the closed-loop system. Event- and self-triggered control designs quite naturally handle this type of delay, and they have proven useful for coordinated autonomous vehicle teams in the event-triggered case [75, 91, 115, 255], and in the self-triggered case [77, 181].

The event- and self-triggered stability problem can be traced back to some initial investigations into sampling strategies for digital controllers [92] as alternatives to the traditional periodic sampling strategy [1, 21]. Using the theory of input-to-state stability [235], P. Tabuada [248] formally examined the time for which a closed-loop system

will remain asymptotically stable in the presence of an outdated sample. From this, one can develop a triggering condition that designates when the controller should be updated. In an event-triggered control implementation, the system state is updated when it deviates from the previous sample by a sufficient amount [1, 21, 104, 172, 194, 202, 248, 281]. In the case of self-triggered control, the decision to update is computed from predictions, based on the last update, of when the system state will surpass a given threshold [17, 18, 144, 162, 163, 275, 282].

Common to most of these works is a deterministic system model, for which sampled data can be used to accurately predict the system state at a future time, thereby defining a path toward an update rule. However, for systems under the influence of disturbances or noise, it may be more difficult to make these predictions or to guarantee that the intended stability results are retained. Along these lines, a few works have examined the robustness of event- or self-triggering frameworks to disturbances. For example, the robustness of a self-triggered control strategy to disturbances was analyzed in [163] for linear systems, and in [42] for parameter uncertainties. In [282], a self-triggered $\mathcal{H}_\infty$ control was developed for linear systems with a state-dependent disturbance, and this was extended in [283] for an exogenous disturbance in $\mathcal{L}_2$ space. The optimality of certainty equivalence for event-triggered LQG systems was examined in [172].

In [3], the authors develop a self-triggering rule for stochastic differential equations (SDEs) that guarantees stability in probability, which, to our knowledge, is the first self-triggering scheme for stochastic control systems alongside [13]. However, there are not

published examples to which the methods in [3] can be applied, perhaps due to the authors' initial assumptions. This is unlike the the deterministic case, for which many examples provide a good sense of the duration between controller updates (see [248] and cf. [194, 281]). Chapter 7 deals with this problem.

Chapter 3

Preliminaries and Road Map

This chapter serves two purposes. First, it introduces some of the preliminary concepts and mathematical tools that will be used in later sections. In doing so, this chapter also assists in drawing a road map of the remainder of this dissertation and motivating the inclusion of the results of Chapters 4-7. Section 3.1 introduces the stochastic modeling approach for describing the motion of autonomous vehicles, including an example of how the relative motion between a vehicle and a target or proximate neighbor can be robustly described by a controlled stochastic differential equation (SDE). Section 3.2 presents some standard results from Lyapunov stability theory that are frequently used to develop stabilizing controllers for the SDEs. In Section 3.3, we examine the problem of minimizing certain costs associated with the SDE solutions, leading to a formulation of a stochastic optimal feedback control problem. This includes an introduction to the Hamilton-Jacobi-Bellman equation, an explanation of how its solution prescribes a stabilizing controller, and a brief summary of the numerical technique that is used in Chapters 4-5 to compute solutions to HJB equations. Section 3.4 includes some of

the issues associated with the numerically-computed stochastic optimal feedback controllers developed in this dissertation, and it also provides comments on a complement to the self-triggered control strategy of Chapter 7.

3.1 Stochastic Modeling Approach

We first introduce some of the models used in this dissertation to describe vehicle motion relative to an unpredictable target or proximate neighbor, although details specific to Chapters 4-6 can be found there. We begin with the Dubins vehicle, which is used to describe UAV motion. Located at position $[x(t),\ y(t)]^{\mathsf{T}}$, the Dubins vehicle moves in the direction of its heading angle θ at a constant speed v. The UAV model is assumed to have sufficient autopilots such that its kinematics may be described by first-order differential equations [198, 205]:

$$\begin{aligned} \mathrm{d}x(t) &= v\cos(\theta)\mathrm{d}t \\ \mathrm{d}y(t) &= v\sin(\theta)\mathrm{d}t \\ \mathrm{d}\theta(t) &= \alpha\left(\theta_c - \theta\right)\mathrm{d}t, \end{aligned}$$

where θ_c is the commanded heading angle, and $\alpha > 0$ is a system parameter. With the choice of $\theta_c = \theta + u/\alpha_\theta$, the UAV can be modeled as a planar Dubins vehicle [82]:

$$\text{Dubins vehicle} \begin{cases} \mathrm{d}x(t) = v\cos\left(\theta(t)\right)\mathrm{d}t \\ \mathrm{d}y(t) = v\sin\left(\theta(t)\right)\mathrm{d}t \\ \mathrm{d}\theta(t) = u\mathrm{d}t, \qquad u \in \mathcal{U} \end{cases} \tag{3.1}$$

where the bounded turning rate $u \in \mathcal{U} \equiv \{u : |u| \leq u_{\max}\}$ must be found. In Chapter 6, we also consider the case where the vehicle can control its acceleration with a second control variable a, which must also be found:

$$\text{Unicycle} \begin{cases} \mathrm{d}x(t) = v(t)\cos(\theta(t))\,\mathrm{d}t \\ \mathrm{d}y(t) = v(t)\sin(\theta(t))\,\mathrm{d}t \\ \mathrm{d}\theta(t) = u\mathrm{d}t \\ \mathrm{d}v(t) = a\mathrm{d}t. \end{cases} \tag{3.2}$$

We refer this vehicle model the unicycle. Due to the approach used in Chapter 6, we do not bound the control inputs u or a for the unicycle, but control inputs will be penalized, as described in Chapter 6, so that they remain small.

As an example to motivate how one can deal with unknown future motion, let the current position of a proximate vehicle or target be (x_T, y_T). If the neighboring vehicle is identical to the tracking vehicle, it could have a model identical to (3.1) or (3.2). However, without high levels of communication between the two vehicles, the control inputs u, or u and a, of the neighboring vehicle are unknown. It could also be the case that the proximate vehicle has completely unknown kinematics. In this case, the evolution of the states $x_T(t)$ and $y_T(t)$ is unknown.

In the problem formulation, we consider the unknown motion as a stochastic process. Drawing from the field of estimation, the simplest model that can be used to describe an unknown signal suggests the use of a Wiener process [108]. In the case of an identical neighboring vehicle, but one with unknown control inputs, we can define

a stochastic Dubins vehicle[1] and stochastic unicycle as the complements to (3.1) and (3.2), respectively.

$$\text{Stochastic Dubins vehicle} \begin{cases} \mathrm{d}x_T(t) = v\cos\left(\theta_T(t)\right)\mathrm{d}t \\ \mathrm{d}y_T(t) = v\sin\left(\theta_T(t)\right)\mathrm{d}t \\ \mathrm{d}\theta_T(t) = \sigma\mathrm{d}w_\theta \end{cases} \tag{3.3}$$

$$\text{Stochastic Unicycle} \begin{cases} \mathrm{d}x_T(t) = v_T(t)\cos\left(\theta_T(t)\right)\mathrm{d}t \\ \mathrm{d}y_T(t) = v_T(t)\sin\left(\theta_T(t)\right)\mathrm{d}t \\ \mathrm{d}\theta_T(t) = \sigma\mathrm{d}w_\theta \\ \mathrm{d}v_T(t) = \sigma_v\mathrm{d}w_v \end{cases} \tag{3.4}$$

Here, $\mathrm{d}w_v$ and $\mathrm{d}w_\theta$ are mutually increments of a Wiener process, and σ is a known noise intensity that characterizes the uncertainty in the kinematics, control, or environmental influences. We will also consider the case where the neighboring vehicle's control inputs are known in expectation, but otherwise uncertain, so that $\mathrm{d}\theta_T = u_T\mathrm{d}t + \sigma_\theta\mathrm{d}w_\theta$, for example. The model (3.3), while formulated here as the motion of a Dubins vehicle with unknown turning rate control, is also used in Chapter 5 to describe wind incident at a stochastically-varying direction. The model (3.4) is used in Chapter 6 to describe vehicle kinematics in a formation control problem. For a neighboring vehicle or other subject that, unlike (3.3)-(3.4), does not have a known heading angle, the target model

[1]Without a bounded turning rate, a more appropriate name for this model might be a stochastic unicycle, but we wish to differentiate the fixed-speed case from the controlled-speed case.

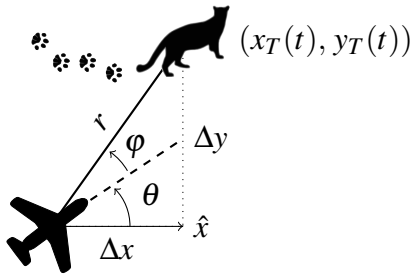


Figure 3.1: Top-down view of a Dubins vehicle at a distance $r(t)$ and viewing angle $\varphi(t)$ to a subject under surveillance with unknown future trajectory.

we use to describe its unknown future motion is

$$\begin{aligned} \mathrm{d}x_T(t) &= \sigma \mathrm{d}w_x \\ \mathrm{d}y_T(t) &= \sigma \mathrm{d}w_y. \end{aligned} \tag{3.5}$$

In addition to describing the motion of an unknown target in Chapter 4, (3.5) is also used in Chapter 5 to model a stochastic wind (without known direction) as it affects the motion of a Dubins vehicle (3.1).

For each of these models, we can derive an equation that describes the relative motion between a Dubins vehicle / unicycle and a proximate vehicle or target. As an example (studied in more detail in Chapter 4), consider the case of a Dubins vehicle (3.1) that should track a target with unknown future motion described by (3.5), as shown in Fig. 3.1. Denoting $\Delta x(t) = x_T(t) - x(t)$ and $\Delta y_T(t) = y_T(t) - y(t)$ as the Cartesian components of distance, the evolution of $r(t) = \sqrt{(\Delta x(t))^2 + (\Delta y(t))^2}$ can be found

using Itô's Lemma [105] as

$$\mathrm{d}r(t) = \frac{\Delta x}{r}\mathrm{d}(\Delta x) + \frac{\Delta y}{r}\mathrm{d}(\Delta y) + \frac{1}{2}\left(\frac{1}{r} - \frac{(\Delta x)^2}{r^3}\right)(\mathrm{d}(\Delta x))^2 + \frac{1}{2}\left(\frac{1}{r} - \frac{(\Delta y)^2}{r^3}\right)(\mathrm{d}(\Delta y))^2 - \frac{(\Delta x)(\Delta y)}{r^3}(\mathrm{d}(\Delta x))(\mathrm{d}(\Delta y)).$$

With substitution of (3.1) and (3.5) and by taking into account that $\cos(\theta+\varphi) = \Delta x/r$ and $\sin(\theta+\varphi) = \Delta y/r$ (Fig. 3.1), it can be shown that

$$\mathrm{d}r(t) \quad = \left(-v_A\cos\varphi + \tfrac{\sigma^2}{2r}\right)\mathrm{d}t + \sigma\cos(\theta+\varphi)\,\mathrm{d}w_x + \sigma\sin(\theta+\varphi)\,\mathrm{d}w_y.$$

Similarly, if we express $\varphi(t) = \tan^{-1}(\Delta y(t)/\Delta x(t)) - \theta(t)$,

$$\begin{aligned}\mathrm{d}\varphi(t) &= -\frac{\Delta y}{r^2}\mathrm{d}(\Delta x) + \frac{\Delta x}{r^2}d(\Delta y) + \frac{(\Delta x)(\Delta y)}{r^4}(\mathrm{d}(\Delta x))^2 - \frac{(\Delta x)(\Delta y)}{r^4}(\mathrm{d}(\Delta y))^2 - u\mathrm{d}t \\ &= \left(\frac{v_A}{r}\sin\varphi - u\right)\mathrm{d}t - \frac{\sigma}{r}\sin(\theta+\varphi)\,\mathrm{d}w_x + \frac{\sigma}{r}\cos(\theta+\varphi)\,\mathrm{d}w_y.\end{aligned}$$

Due to the invariance of a Wiener process to a rotation of the coordinate system [105], the evolution of $r(t)$ and $\varphi(t)$ simplify as

$$\begin{aligned}\mathrm{d}r(t) &= \left(-v\cos\varphi + \frac{\sigma^2}{2r}\right)\mathrm{d}t + \sigma\mathrm{d}w_0 \\ \mathrm{d}\varphi(t) &= \left(\frac{v}{r} - u\sin\varphi\right)\mathrm{d}t + \sigma\mathrm{d}w_\perp.\end{aligned}$$

Denoting $\mathbf{x}(t) = [r(t), \varphi(t)]^\intercal$ we can write down a stochastic differential equation (SDE) [157] describing the relative motion between the DV and its target as

$$\mathrm{d}\mathbf{x}(t) = f(\mathbf{x},u)\mathrm{d}t + g(\mathbf{x})\mathrm{d}\mathbf{w}, \tag{3.6}$$

where $f(\mathbf{x},u)$ describes the drift of the relative state between the tracking agent and its target, and where $g(\mathbf{x})$ captures the known noise intensity of the stochastic motion. As before, d$\mathbf{w}$ are increments of a Wiener process. Similarly to this example, it is not

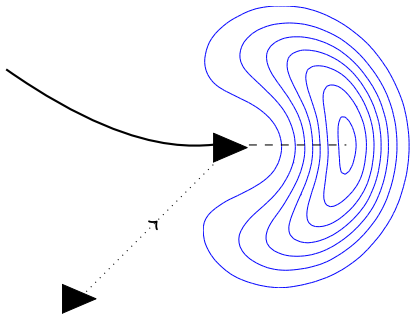


Figure 3.2: Illustration of a probability density function of a non-deterministic target's future position (computed from stochastic simulations of (3.4)). The tracking agent should respond to both the target's current relative state and the distribution of its future motion.

difficult to derive similar SDEs for other autonomous vehicle and target models. The stochastic process $\mathbf{x}(t)$ is formally defined on a complete probability space $(\Omega, P, \mathcal{F})$ comprised of a set of possible outcomes Ω, a family $\mathcal{F}$ of subsets of Ω, and a probability measure $P : \mathcal{F} \to [0,1]$ [157].

Under certain conditions on $f(\cdot)$ and $g(\cdot)$, the SDE (3.6) admits a solution $\mathbf{x}(t)$, and we offer two interpretations of $\mathbf{x}(t)$ as it pertains the the problems presented in this dissertation. Fixing an element $\omega \in \Omega$, the solution $\mathbf{x}_\omega(t)$ describes a sample path of the relative vehicle-target state as a function of time, one out of many drawn from a family of trajectories. In this sense, we account for a wide variety of possible influences from the environment, neighboring vehicle control algorithms, and kinematic uncertainties. However, as the sample paths of the Wiener process are almost never differentiable [183], it is probably most appropriate to say that this SDE sample path $\mathbf{x}(t)$ is "close to" a possible outcome of the state caused by a real vehicle trajectory. In an alternative viewpoint, we can fix $t \in [0,T]$, in which case $\mathbf{x}_\omega(t) \in \mathbb{R}^n$ is a random

variable. The stochastic model therefore induces a prior probability distribution that describes the probability of finding the relative system state in an interval $(\mathbf{x}, \mathbf{x}+d\mathbf{x})$ at a particular future time [280]. For example, the distribution for the model (3.4) is the so-called banana distribution [154] and is illustrated (simulated) in Fig. 3.2. In solving a control problem based on these SDEs, we not only compute the optimal control with respect to the current system state, but also with respect to the distribution of all possible trajectories originating from the current state, leading to the notion that the control *anticipates* the uncertainty inherent to the system.

Introducing a stochastic process to replace unknown dynamics or kinematics is not a new idea. Surveys of stochastic state space models for uncertain maneuvering in target tracking can be found in [146] and in [66] for models in biology, such as the movement of micro-organisms and larger animals. Some previous work has also examined stochasticity as a replacement for known, fast time-scale, deterministic motion [123, 129]. Control problems involving these stochastic models, especially in autonomous vehicle control, are much less common.

In the aforementioned models, we have assumed that the target is continuously observed, and any noise due to observation error is assumed to be included in the noise intensity of the model, i.e., the parameter σ. However, it may be the case that observation of the target is occluded by obstacles or sensory interference, and this possibility is tackled in Chapter 4.

3.2 Stochastic Stability and Stabilization via the Lyapunov Direct Method

Stochastic stability is concerned with the qualitative behaviors of the solutions to SDEs like (3.6), including regularity, boundedness, and the tendency for the solution $\mathbf{x}(t)$ to asymptotically reach the origin or a neighborhood around it. There is an extensive list of notions of stochastic stability [260], including almost sure stability [131], moment stability [157], and stochastic input-to-state stability [152], to name a few. However, in most of the problems tackled in this dissertation, the noise scaling factor $g(\mathbf{x})$ in (3.6) does not necessarily vanish at the origin (e.g., $g(0) \neq 0$). For example, in (3.5), the stochastic motion of the target is always present, and we should not expect the origin $\mathbf{x} = 0$ to admit an equilibrium point or trivial solution $\mathbf{x}(t) \equiv 0$. However, the assumption $g(0) = 0$ is a common assumption in many stochastic stability results, including those in [131, 157]. For that reason, we first focus here on a notion of disturbance attenuation explored in [135].

The differential operator $\mathcal{L}^u$ associated with the system (3.6) under a control u, when applied to a positive-definite, twice differentiable function $V : \mathbb{R}^n \to \mathbb{R}^+$, is

$$\mathcal{L}^u V(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} f_i(\mathbf{x}, u) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 V}{\partial x_i \partial x_j} g_{ij}(\mathbf{x}). \tag{3.7}$$

Define a positive constant c, strictly increasing functions α_1 and α_2 with the properties that $\alpha_i(0) = 0$ and $\alpha_i(r) \to \infty$ as $r \to \infty$, $i = 1, 2$ (i.e., they are of class $\mathcal{K}_\infty$ [130]), and an increasing function $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$. Assume for now that the diffusion scaling factor

$g(\cdot)$ is a constant and independent of $\mathbf{x}$. If, for all $\mathbf{x} \in \mathbb{R}^n$, $V(\mathbf{x})$ is such that

$$\alpha_1(|\mathbf{x}|) \leq V(\mathbf{x}) \leq \alpha_2(|\mathbf{x}|) \tag{3.8}$$

$$\mathcal{L}^u V(\mathbf{x}) \leq -cV(\mathbf{x}) + \gamma(|gg^\intercal|), \tag{3.9}$$

then the expected value of $V(\mathbf{x})$ satisfies [135]

$$\mathbb{E}(V(\mathbf{x})) \leq e^{-ct} V(\mathbf{x}(0)) + c^{-1}\gamma(|gg^\intercal|), \tag{3.10}$$

so that $\lim_{t\to\infty} \mathbb{E}V(\mathbf{x}(t)) \leq \gamma(|gg^\intercal|)/c$.

In the case of a state-dependent $g(\cdot)$ such that $g(0) = 0$, we can also discuss stability in probability [157, Section 4.2]. Suppose the function positive-definite, twice-differentiable function $V(\mathbf{x})$ is such that $\mathcal{L}^u V \leq 0$ for all $\mathbf{x}$. Then for every pair $\varepsilon \in (0,1)$ and $r > 0$, there exists a $\delta(\varepsilon) > 0$ such that

$$P(|\mathbf{x}| < r \text{ for all } t \geq 0) \geq 1 - \varepsilon$$

whenever $|\mathbf{x}(0)| < \delta(\varepsilon)$. If, alternatively, we had a Lyapunov function satisfying

$$\mathcal{L}^u V(\mathbf{x}) \leq -k(\mathbf{x}) \leq 0 \tag{3.11}$$

for some positive-definite function $k(\mathbf{x})$, then the state $\mathbf{x}(t)$ is stochastically asymptotically stable, that is, for every $\varepsilon \in (0,1)$ there exists a $\delta(\varepsilon) > 0$ such that

$$P\left(\lim_{t\to\infty} \mathbf{x}(t) = 0\right) \geq 1 - \varepsilon.$$

We will call the function $V(\mathbf{x})$ a stochastic Lyapunov function (or Lyapunov function for short) if it satisfies one of these types of conditions.

The Lyapunov direct method involves the analytic construction of a Lyapunov func-

tion $V(\mathbf{x})$ that satisfies the desired conditions for stability, e.g., stability in probability. Oftentimes, both the Lyapunov function and control $u(\mathbf{x})$ are constructed simultaneously. If this process is successful, the resulting control is a (stochastically) stabilizing control law with respect to the control Lyapunov function (CLF) $V(\mathbf{x})$. In this case, the system is called feedback-stabilizable.

The analytic form of the control law and its Lyapunov function can be quite appealing. In particular, provable claims of stability, safety, or performance may be derived from the analytic form of $\mathcal{L}^u V(\mathbf{x})$. However, when kinematic nonlinearities and stochasticity are included, the control design problem becomes significantly more complex. The ability to find a stabilizing controller often boils down to trial and error. In some cases, recipe-like techniques, such as (stochastic) backstepping [135], can be followed to compute a stabilizing controller. Coefficients of parameterized Lyapunov functions also can sometimes be found using direct optimization [187, 188].

Next, let us suppose that the differential operator $\mathcal{L}^u$ is such that (3.11) holds with equality, i.e.,

$$\mathcal{L}^u V(\mathbf{x}) + k(\mathbf{x}) = 0. \tag{3.12}$$

Equation (3.12) is a partial differential equation[2]. As we will see in the following section, the solution $V(\mathbf{x})$ to this PDE is the expected total cost incurred by the trajectory $\mathbf{x}(t)$ under the control $u(\mathbf{x})$ as it accumulates the instantaneous cost $k(\mathbf{x})$ over an infinite horizon. In other words, under some additional conditions on the $\mathcal{L}^u$, $V(\mathbf{x})$, and $k(\mathbf{x})$,

[2]Boundary conditions are considered in the following section.

(3.12) has solution [131, Chapter 8]

$$V(\mathbf{x}) = \mathbb{E}\left(\int_0^\infty k(\mathbf{x}(t))\mathrm{d}t\right). \tag{3.13}$$

In the context of the CLF approach, this means that when constructing a CLF, there is an implicit cost associated with the state $\mathbf{x}(t)$ under the control $u(\mathbf{x})$, but this cost is just that which appears on the right hand side of (3.11) during the construction of the CLF. It seems perhaps more natural and useful to instead begin with a desired objective $k(\mathbf{x})$ and, from there, develop a CLF that achieves a good performance with respect to this objective. This idea has previously been explored in [99, 100, 197], for example. Notably, if the feedback control $u(\mathbf{x})$ is that which minimizes $V(\mathbf{x})$ in (3.13), it is an optimal feedback control, which brings us to the topic of the next section.

3.3 Stochastic Optimal Feedback Control

We now state two types of cost functionals for (3.6) that will be used to define optimal feedback control problems.

$$\textbf{(P1)} \qquad J(\mathbf{x},u) = \mathbb{E}\left\{\int_0^\infty e^{-\beta t}k(\mathbf{x}(t),u)dt\right\}$$

$$\textbf{(P2)} \qquad J(\mathbf{x},u) = \mathbb{E}\left\{h(\mathbf{x}(\tau)) + \int_0^\tau k(\mathbf{x}(t),u)dt\right\}$$

The first (P1) is an infinite-horizon, discounted cost function that incurs instantaneous cost $k(\mathbf{x},u)$, but with future costs weighted less heavily than immediate costs. In this dissertation, we employ the discounted, infinite-horizon approach for the problem of maintaining (over an infinite-horizon) a nominal distance from a target with an un-

known future trajectory. The second (P2) is a stochastic shortest-path problem [46], sometimes framed as "control until a target set is reached" [137], in which cost $k(\mathbf{x}, u)$ is incurred only until the time $\tau = \inf\{t > 0 : \mathbf{x} \notin G^0\}$, the first exit of the state $\mathbf{x}(t)$ from the interior $G^0 = \text{int}(G)$ of the set G, at which point a terminal cost $h(\mathbf{x})$ is added (for minimum-time problems, $k(\mathbf{x}, u) = 1$ and $h(\mathbf{x}) = 0$ for all $\mathbf{x}, u$). We use the shortest-path approach for the problem of navigating a Dubins vehicle in minimum time toward a waypoint or target in a stochastic wind.

3.3.1 The Hamilton-Jacobi-Bellman Equation

In each problem (P1) and (P2), $J(\mathbf{x}, u)$ is the total cost under the control mapping u. Let $V(\mathbf{x})$ be the infimum over all admissible controls[3] $u \in \mathcal{A}$:

$$V(\mathbf{x}) = \inf_{u \in \mathcal{A}} J(\mathbf{x}, u)$$

and is called the value of the cost-to-go function (or value function). The HJB PDE for the discounted control problem for $\mathbf{x} \in G^0 \subset \mathbb{R}^n$ is

$$0 = \inf_{u \in \mathcal{A}} \{-\beta V(\mathbf{x}) + \mathcal{L}^u V(\mathbf{x}) + k(\mathbf{x}, u)\} \tag{3.14}$$

and, for $\mathbf{x} \in \partial G$, the boundary of G, we have the boundary condition

$$\nabla V^T(\mathbf{x}) r(\mathbf{x}) = 0 \qquad \text{for } \mathbf{x} \in \partial G \text{ (reflective)} \tag{3.15}$$

[3] A control $U_t \equiv u(\mathbf{x}(t))$ is considered admissible here if a) it is non-anticipative with respect to the Wiener process $w(t)$ (i.e., U_t is $\{\mathcal{F}_t\}$-adapted [183])" b) it satisfies any given control constraints; c) the SDE (3.6) has a strong solution; d) given a twice differentiable function $\phi(\mathbf{x})$, it ensures that $\int_0^t e^{-\beta s} \nabla\phi(\mathbf{x}(s)) g(\mathbf{x}(s)) d\mathbf{w}_s$ is a martingale [183], and e) $\lim_{t\to\infty} e^{-\beta t} \mathbb{E} V(\mathbf{x}(t)) = 0$

where $r(\mathbf{x})$ is the reflection direction. For the shortest-path problem, the HJB equation reads

$$0 = \inf_{u \in \mathcal{A}} \{\mathcal{L}^u V(\mathbf{x}) + k(\mathbf{x}, u)\} \tag{3.16}$$

with boundary conditions

$$\begin{aligned} V\mathbf{x}) &= h(\mathbf{x}) \qquad \text{for } \mathbf{x} \in \partial G \text{ (absorbing)} \\ \nabla V^T(\mathbf{x}) r(\mathbf{x}) &= 0 \qquad \text{for } \mathbf{x} \in \partial G \text{ (reflective)} \end{aligned}$$

where $h(\mathbf{x})$ is a cost incurred when the process $\mathbf{x}(t)$ first exists G^0. In the context of the problems studied in this thesis, the solution to the HJB, the value function, measures the total cost for a vehicle to implement its optimal control over the planning horizon $[0, \infty)$ or $[0, \tau]$ defined in (P1) and (P2), respectively. This optimal input is given by

$$u^* = \arg\min_{u \in \mathcal{A}} \{-\beta V(\mathbf{x}) + \mathcal{L}^u V(\mathbf{x}) + k(\mathbf{x}, u)\}$$

for (P1), and for (P2), it reads

$$u^* = \arg\min_{u \in \mathcal{A}} \{\mathcal{L}^u V(\mathbf{x}) + k(\mathbf{x}, u)\}.$$

Rather than showing that the PDEs (3.14) and (3.16) can be derived from their respective optimal control problems under costs (P1) and (P2), which can involve tricky claims of smoothness[4], we approach this issue from the opposite direction and verify that a given solution to the HJB equation is the value of the optimal control problem cost-to-go. Furthermore, we show that the control provided by that HJB solution is the

[4] In particular, the question of whether or not a value function is sufficiently smooth to satisfy the HJB equation led to the theory of viscosity solutions to the HJB equations [36], which are beyond the scope of this dissertation.

optimal control, i.e., the total cost of using the optimal control is equal to the value function. We show this for the discounted, infinite-horizon control problem, and the shortest-path problem follows similarly with $\beta = 0$.

Assume that we have computed a twice-differentiable solution $\phi(\mathbf{x})$ to the HJB equation (3.14)-(3.15), and let u^* be the unique value of $u \in \mathcal{A}$ that minimizes $-\beta\phi(\mathbf{x}) + \mathcal{L}^u\phi(\mathbf{x}) + k(\mathbf{x},u)$:

$$0 = \inf_{u\in\mathcal{A}} \{-\beta\phi(\mathbf{x}) + \mathcal{L}^u\phi(\mathbf{x}) + k(\mathbf{x},u)\}$$

$$u^*(\mathbf{x}) = \arg\min_{u\in\mathcal{A}} \{-\beta\phi(\mathbf{x}) + \mathcal{L}^u\phi(\mathbf{x}) + k(\mathbf{x},u)\}.$$

Let $\mathbf{x}^*$ be the corresponding state process with initial condition $\mathbf{x}$ under the control $u^*(\mathbf{x}(t))$. Applying Itô's rule to the function $e^{-\beta t}\phi(\mathbf{x}^*)$, we have for $t \in [0,\infty)$

$$\begin{aligned} e^{-\beta t}\phi(\mathbf{x}^*) = \phi(\mathbf{x}) &+ \int_0^t e^{-\beta s}\left(-\beta\phi(\mathbf{x}^*(s)) + \mathcal{L}^{u^*}\phi(\mathbf{x}^*(s))\right)\mathrm{d}s \\ &+ \int_0^t e^{-\beta s} g(\mathbf{x}^*(s))^\intercal \nabla\phi(\mathbf{x}^*(s))\mathrm{d}\mathbf{w}_s. \end{aligned}$$

Adding $\int_0^t k(\mathbf{x}^*(s), u^*(\mathbf{x}^*(s)))\mathrm{d}s$ and taking the expectation of both sides,

$$\begin{aligned} &\mathbb{E}\int_0^t e^{-\beta s}k(\mathbf{x}^*(s),u^*(\mathbf{x}^*(s)))\mathrm{d}s + e^{-\beta t}\mathbb{E}(\phi(\mathbf{x}^*(t))) \\ &\quad = \phi(\mathbf{x}) + \mathbb{E}\int_0^t e^{-\beta s}\left(-\beta\phi(\mathbf{x}^*(s)) + \mathcal{L}^{u^*}\phi(\mathbf{x}^*(s)) + k(\mathbf{x}^*(s),u^*(\mathbf{x}^*(s)))\right)\mathrm{d}s. \end{aligned}$$

From the HJB equation (3.14), the right hand side integrand is zero. Letting $t \to \infty$,

$$\phi(\mathbf{x}) = \mathbb{E}\int_0^\infty e^{-\beta s}k(\mathbf{x}^*(s),u^*(\mathbf{x}^*(s)))\mathrm{d}s = J(\mathbf{x},u^*). \tag{3.17}$$

We can now repeat the above derivation for an arbitrary $u \in \mathcal{A}$. In doing so, instead of the integrand vanishing due to the HJB equation, we now have $-\beta\phi(\mathbf{x}(s)) +$

$\mathcal{L}^u\phi(\mathbf{x}(s)) + k(\mathbf{x}(s),u) \geq 0$, and so

$$\phi(\mathbf{x}) \leq \mathbb{E}\int_0^\infty e^{-\beta s}k(\mathbf{x}(s),u)\mathrm{d}s = J(\mathbf{x},u).$$

Since $u \in \mathcal{A}$ is arbitrary, we can say $\phi(\mathbf{x}) \leq \inf_{u\in\mathcal{A}} J(x,u) = V(\mathbf{x})$, but from (3.17), $\phi(\mathbf{x}) = J(\mathbf{x},u^*)$, and so we have that $V(\mathbf{x}) = J(\mathbf{x},u^*)$, implying that u^* is the optimal control.

3.3.2 As a Stabilizing Controller

Part of our motivation for taking a stochastic optimal feedback control approach revolves around the fact that the difficulty of finding a stabilizing Lyapunov function-based controller (as described in Section 3.2) can be avoided by instead solving the HJB equation. This is because the solution to the HJB, the value function, prescribes a control that guarantees stability of the trajectory in state space. In this sense, the value function serves as a Lyapunov function that can be constructed through the solution of a PDE. To see how the value function provides stability for the discounted problem, we can examine the expected change in the cost-to-go in a manner analogous to [131]. Applying Itô's lemma to the function $e^{-\beta t}V(\mathbf{x})$:

$$\begin{aligned}
\mathrm{d}\left(e^{-\beta t}V(\mathbf{x})\right) &= e^{-\beta t}\left(-\beta V(\mathbf{x})\mathrm{d}t + \mathrm{d}x^\intercal\nabla V(\mathbf{x}) + \frac{1}{2}\mathrm{d}\mathbf{x}^\intercal\nabla^2 V(\mathbf{x})\mathrm{d}\mathbf{x}\right) \\
&= e^{-\beta t}\left(-\beta V(\mathbf{x}) + f(\mathbf{x},u)^T\nabla V + \frac{1}{2}g(\mathbf{x})^T\nabla^2 V(\mathbf{x})g(\mathbf{x})\right)\mathrm{d}t \\
&\quad + e^{-\beta t}g(\mathbf{x},t)^\intercal V(x)\mathrm{d}\mathbf{w}. \qquad (3.18)
\end{aligned}$$

Taking the expected value of both sides and adding and subtracting $e^{-\beta t}k(\mathbf{x},u)$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-\beta t}\mathbb{E}(V(\mathbf{x}))\right)$$
$$= e^{-\beta t}\mathbb{E}\left(-\beta V(\mathbf{x}) + f(\mathbf{x},u)^T\nabla V + \frac{1}{2}g(\mathbf{x})^T\nabla^2 V(\mathbf{x})g(\mathbf{x}) + k(\mathbf{x},u) - k(\mathbf{x},u)\right)$$

Inserting the HJB equation with the minimizing control (3.14) to cancel terms, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-\beta t}\mathbb{E}(V(\mathbf{x}))\right) = -e^{-\beta t}k(\mathbf{x},u) \leq 0.$$

As this quantity is decreasing, we can then show using Markov's inequality that $e^{-\beta t}V(\mathbf{x}(t))$ will asymptotically reach zero with probability one. Fix some constant $c > 0$. Then

$$\lim_{t\to\infty} P\left(e^{-\beta t}V(\mathbf{x}(t)) \geq c\right) \leq \frac{1}{c}\lim_{t\to\infty}\mathbb{E}(e^{-\beta t}V(\mathbf{x}(t))) = 0.$$

3.3.3 The Markov Chain Approximation Method

While there are many numerical methods to compute an optimal feedback control from the HJB equation in literature (see, for example, [112]), we employ the Markov chain approximation method [137] for numerically determining the optimal feedback control corresponding to the controlled diffusion process (3.6) and one of the HJB equations (3.14) or (3.16). Since we use this method for computing the vehicle control policies in Chapters 4-6, we now give a brief overview of the method.

When discretizing a state space for dynamic programming in stochastic problems, it is often the case that the chosen spatial and temporal step sizes do not accurately scale in the same way as the stochastic process. The Markov chain approximation

method explicitly ensures this notion of consistency of the discretized process with the underlying diffusion process, and is a well-accepted technique in stochastic control (although it is perhaps used infrequently in autonomous vehicle applications).

The technique involves the construction of a Markov chain $\left\{\xi_n^h,\ n<\infty\right\}$ on a discretized state space with transition probabilities $p^h(\mathbf{y}|\mathbf{x},u)$ from state $\mathbf{x}\in\mathbb{R}^n$ to state $\mathbf{y}\in\mathbb{R}^n$ under an admissible control $u\in\mathcal{A}$. The transition probabilities can be found from the coefficients in the finite-difference approximations to the operator $\mathcal{L}^u$ in (3.7). Assuming for now that the noise is uncorrelated, i.e., $\{g(\mathbf{x})g(\mathbf{x})^\intercal\}_{ij}=0$ for $i\neq j$, and that the discretized state space $S_h=\{\mathbf{x}\,:\,\mathbf{x}=h\sum_i \mathbf{e}_i m_i,\ m_i=0,\pm1,\pm2,\ldots\}$ is a uniform grid on bases $\mathbf{e}_i$, $i=1,\ldots,n$, we can apply up-wind approximations for first-order derivatives and the standard approximation for second-order derivatives. Then since $J(\mathbf{x},u)$ satisfies, for fixed u,

$$0=-\beta J(\mathbf{x},u)+\mathcal{L}^u J(\mathbf{x},u)+k(\mathbf{x},u),$$

a finite-difference discretization for $J(\mathbf{x},u)$ with step size h is

$$0=\sum_{i=1}^n (f_i(\mathbf{x},u))^+\left[\frac{J(\mathbf{x}+\mathbf{e}_i h,u)-J(\mathbf{x},u)}{h}\right]+\sum_{i=1}^n (f_i(\mathbf{x},u))^-\left[\frac{J(\mathbf{x},u)-J(\mathbf{x}-\mathbf{e}_i h,u)}{h}\right]$$
$$+\frac{1}{2}\sum_{i=1}^n g_i(\mathbf{x})^2\left[\frac{g(\mathbf{x}+\mathbf{e}_i h)-2g(\mathbf{x})+g(\mathbf{x}-\mathbf{e}_i h)}{h^2}\right]-\beta J(\mathbf{x},u)+k(\mathbf{x},u)$$

where $(a)^+=\max\{a,0\}$ and $(a)^-=\max\{-a,0\}$. Collecting the $J(\mathbf{x},u)$ terms on one

side, we can identify a finite difference form for $J(\mathbf{x},u)$ as

$$\begin{aligned} J^h(\mathbf{x},u) &= \frac{\sum_{i=1}^{n}\left(\frac{|f_i(\mathbf{x},u)|}{h}+\frac{g_i(\mathbf{x})^2}{h^2}\right)}{\beta+\sum_{i=1}^{n}\left(\frac{|f_i(\mathbf{x},u)|}{h}+\frac{g_i(\mathbf{x})^2}{h^2}\right)}\sum_y p^h(\mathbf{y}|\mathbf{x},u)J^h(\mathbf{y},u)+k(\mathbf{x},u)\Delta t^h(\mathbf{x},u) \\ &= \left(1-\beta\Delta t^h(\mathbf{x},u)\right)\sum_y p^h(\mathbf{y}|\mathbf{x},u)J^h(\mathbf{y},u)+k(\mathbf{x},u)\Delta t^h(x,u) \\ &\approx \exp\left\{-\beta\Delta t^h(\mathbf{x},u)\right\}\sum_y p^h(\mathbf{y}|\mathbf{x},u)J^h(\mathbf{y},u)+k(\mathbf{x},u)\Delta t^h(\mathbf{x},u) \end{aligned} \tag{3.19}$$

where

$$p^h(\mathbf{x}\pm\mathbf{e}_i h|\mathbf{x},u)=\frac{\left(\frac{(f_i(\mathbf{x},u))^{\pm}}{h}+\frac{g_i(\mathbf{x})^2}{2h^2}\right)}{\sum_{i=1}^{n}\left(\frac{|f_i(\mathbf{x},u)|}{h}+\frac{g_i(\mathbf{x})^2}{h^2}\right)}$$

are interpreted as the transition probabilities for the Markov chain ξ_n^h, and

$$\Delta t(\mathbf{x},u)=\left[\beta+\sum_{i=1}^{n}\left(\frac{|f_i(\mathbf{x},u)|}{h}+\frac{g_i(\mathbf{x})^2}{h^2}\right)\right]^{-1}$$

is an interpolation interval for the chain. This interpolation interval allows us to define piece-wise constant, continuous parameter interpolations $\xi^h(t)$ and $u^h(t)$ by

$$\xi^h(t)=\xi_n^h,\qquad u^h(t)=u_n^h,\qquad t\in\left[t_n^h,t_n^h+\Delta t^h(x,u(x))\right)$$

The net effect of this representation is a Markov chain whose increments $\Delta\xi_n^h=\xi_{n+1}^h-\xi_n^h$ are statistically close to the original controlled diffusion. Specifically, the chain obeys the following property of local consistency

$$\begin{aligned} &\mathbb{E}\left[\Delta\xi_n^h\right]=f(\mathbf{x},u)\Delta t_n^h(\mathbf{x},u)+o\left(\Delta t_n^h(\mathbf{x},u)\right) \\ &\mathbb{E}\left[\left(\Delta\xi_n^h-\mathbb{E}\Delta\xi_n^h\right)\left(\Delta\xi_n^h-\mathbb{E}\Delta\xi_n^h\right)^{\mathsf{T}}\right]=g(\mathbf{x})g(\mathbf{x})^{\mathsf{T}}\Delta t_n^h(\mathbf{x},u)+o\left(\Delta t_n^h(\mathbf{x},u)\right) \\ &\lim_{h\to 0}\sup_n\left|\xi_{n+1}^h-\xi_n^h\right|=0. \end{aligned}$$

The recursive equation for dynamic programming on the optimal cost for the Markov

chain ξ_n^h follows from taking the infimum over all admissible control sequences in (3.19) and using the principle of optimality:

$$V^h(\mathbf{x}) = \inf_{u \in \mathcal{A}} \left\{ \sum_y e^{-\beta \Delta t(\mathbf{x},u)} p^h(\mathbf{y}|\mathbf{x},u) V^h(\mathbf{y}) + k(\mathbf{x},u) \Delta t^h(\mathbf{x},u) \right\} \qquad \mathbf{x} \in G^h$$

Boundary conditions for absorbing states $\mathbf{x} \in \partial G^h$ are $V^h(\mathbf{x}) = h(\mathbf{x})$. Boundary conditions for reflective states should be chosen to that the Markov chain is locally consistent to not only the process $\mathbf{x}(t)$, but also to the reflection directions $r(\mathbf{x})$. For a rectangular domain with $r(\mathbf{x})$ aligned with inward-pointing domain boundary normals $\hat{n}_i$, for example, this is

$$V^h(\mathbf{x}) = V^h(\mathbf{x} - \hat{n}_i h) \qquad \mathbf{x} \in \partial G^h.$$

The method of value iterations is used with these recursive equations to compute the control laws in this dissertation. At this point, it might be reasonable to assume that for small h, the value $V^h(\mathbf{x})$ under the optimal control for the chain $\left\{ \xi_n^h,\, n < \infty \right\}$ is close to the optimal value $V(\mathbf{x})$ for the original process. This turns out to be the case, but the relevant proofs [137, Chapters 9-10] are beyond the scope of this work.

3.4 Brief Comments on Controller Implementation

The implementation of the numerical control laws, as opposed to analytic control laws, on autonomous vehicles introduces a few hurdles that we mention briefly here. First, based on our approach, once a control has been computed off-line, it is valid for the parameter regime (e.g., σ, v) used during the computations. While there is likely a

margin for which a given controller can be applied to other parameter ranges, we have not examined the extent to which this is possible, and, for now, the control must be computed off-line for each given set of parameters. Re-scaling of optimal control laws has previously been performed in [95], but only for deterministic kinematic models.

A second issue pertains to the change in the control law from one cell in the discretized state space control law to that of an adjacent cell. Given the real, continuous state of the vehicle, any implementation will have to decide if the control to be applied should be chosen from the nearest cell in the discretized state space, or if it should be interpolated from nearby cells. The former will introduce jumps in the control from one cell to the next, for which the inner vehicle control loop should be capable of handling in reasonable time, while the latter approach could change the performance of the vehicle.

Additionally, the storage requirements for the control law scale with the resolution of the discretized grid and with the number of state space variables. Note that the state space of the formation control problem in Chapter 6, for example, is too large for a feedback control to be stored off-line, but, as will be described later, this method is intended to run in real-time.

One interesting implementation issue concerns the difference between the time scale of the vehicle control loop and that of the actual vehicle dynamics. If a target is moving very quickly, it may move through the discretized state space much faster than the vehicle is capable of updating its control. If this happens, can the vehicle still use

the control based on an outdated observation, and if so, for how long? Conversely, how fast does a vehicle control loop have to be to handle the target dynamics? These types of questions are naturally handled by self-triggered and event-triggered control. At a minimum, one should ensure that the particulars of the implementation do not destroy the stability provided by the HJB equation. In Chapter 7, we propose a self-triggered control scheme for stochastic control systems with analytical form. For completeness, we also comment on a potential self-triggering control update rule for the numerical optimal control laws computed in this thesis, but further study is left for future work.

The following proposed method to develop a self-triggering update rule from the computed cost-to-go $V(\mathbf{x})$ and control $u(\mathbf{x})$ is for the case of a positive discounting factor $\beta > 0$, but the stochastic shortest path problem can be solved similarly by setting $\beta = 0$. The idea rests upon tallying the interpolation intervals defined in Section 3.3.3. Let us again consider the Itô derivative of the function $V(\mathbf{x},t) = e^{-\beta t}V(\mathbf{x}(t))$ under a fixed control $u(\mathbf{x}(t))$ found in (3.18). Using finite difference methods, we can construct a Markov chain $\left\{\tilde{\xi}_n^h, n < \infty\right\}$ on the same state space G^h as the Markov chain ξ_n^h in the previous section, using the same techniques as before. This provides a corresponding interpolation interval

$$\begin{aligned}\widetilde{\Delta t}_n^h &= \widetilde{\Delta t}(\tilde{\xi}_n^h, u_n^h) \\ &= \left[\left(\frac{|-\beta V(\mathbf{x},t) + f(\mathbf{x},u)^T\nabla V + \frac{1}{2}g(\mathbf{x})^T\nabla^2 V(\mathbf{x},t)g(\mathbf{x})|}{h} + \frac{g(\mathbf{x})\nabla V \nabla V^\intercal g(\mathbf{x})^\intercal}{h^2}\right)\right]^{-1}\end{aligned}$$

and leads to the continuous parameter interpolations $\tilde{\xi}(t)$ and $u^h(t)$ defined by

$$\tilde{\xi}(t) = \tilde{\xi}_n^h, \;\; u^h(t) = u_n^h, \; t \in \left[t_n^h, t_n^h + \widetilde{\Delta t}_n^h\right). \tag{3.20}$$

Then the interpolated process $\tilde{\xi}(t)$ (3.20) is locally consistent with the quantity $e^{-\beta t}V(\mathbf{x})$ (3.18) whose stability properties are of interest, that is,

$$\begin{aligned}
&\mathbb{E}_{\mathbf{x},n}^{h,u}\left[\Delta\tilde{\xi}_n^h\right] \\
&\quad = \left(-\beta V(\mathbf{x},t) + f(\mathbf{x},u)^T\nabla V + \frac{1}{2}g(\mathbf{x})^T\nabla^2 V(\mathbf{x},t)g(\mathbf{x})\right)\widetilde{\Delta t}^h(\mathbf{x},u) + o\left(\widetilde{\Delta t}^h(\mathbf{x},u)\right) \\
&\mathbb{E}_{\mathbf{x},n}^{h,u}\left[\left(\Delta\tilde{\xi}_n^h - \mathbb{E}\Delta\tilde{\xi}_n^h\right)\left(\Delta\tilde{\xi}_n^h - \mathbb{E}\Delta\tilde{\xi}_n^h\right)^\intercal\right] = g(\mathbf{x})\nabla V\nabla V^\intercal g(\mathbf{x})^\intercal\widetilde{\Delta t}^h(\mathbf{x},u) + o\left(\widetilde{\Delta t}^h(\mathbf{x},u)\right).
\end{aligned}$$

Using this process, we can now examine the stability of the state $\mathbf{x}(t)$ under sample-and-hold update scheme. Assume that the state has last been sampled (and interpolated to the discretized state space) as $\bar{\mathbf{x}}$, which provides an optimal control $\bar{u}$. We are interested in how long the system will remain stable under the control $\bar{u}$. We create a directed graph $\mathcal{G}$ from the discretized state space G^h as shown in Fig. 3.3 with the following properties

- Each cell in the discretized state space G^h is a node in the graph $\mathcal{G}$
- Any two adjacent cells in the discretized state space are connected with an edge in the graph. The edge length from a cell $\mathbf{x}_i$ to $\mathbf{x}_j$ is given by $\widetilde{\Delta t}(\mathbf{x}_i,\bar{u})$, and note that this quantity is the amount of time that the Markov chain $\tilde{\xi}_n^h$ spends in $\mathbf{x}_i$
- An additional node, denoted †, is added to $\mathcal{G}$. We identify all cells $\mathbf{x}_i$ in the

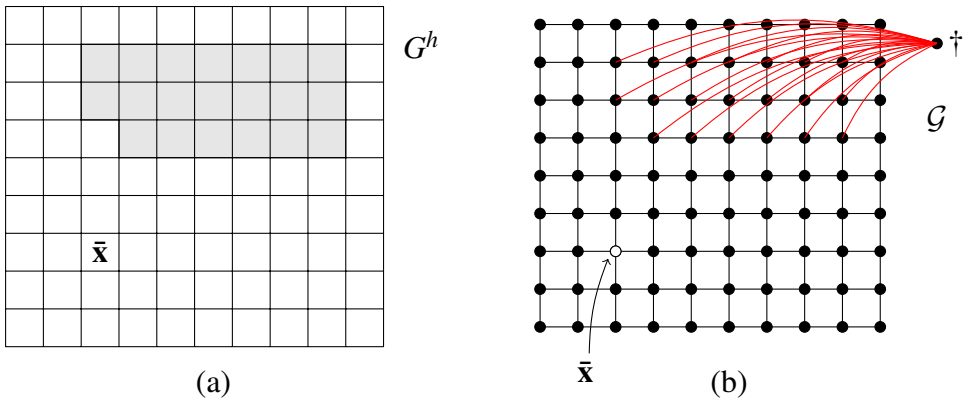


Figure 3.3: (a) The last-observed state $\bar{\mathbf{x}}$ interpolated to the discretized state space G^h. The regions with the property (3.21) are shown in gray, and boundaries ∂G^h are not shown. (b) The corresponding graph $\mathcal{G}$, based on G^h with an added node †. The lengths of the red edges are set as zero, while the black edges' lengths are based the interpolation intervals (see text).

discretized state space that satisfy

$$(f(\mathbf{x}_i, \bar{u}) - f(\mathbf{x}_i, u(x_i)))^\mathsf{T} \nabla V(\mathbf{x}_i) \geq k(\mathbf{x}_i, u(\mathbf{x}_i)) \tag{3.21}$$

and from each of these cells, an edge is added from $\mathbf{x}_i$ to † with edge length 0. (In dealing instead with the stochastic shortest path problem, an edge of length 0 can be added from any terminal cell to †.)

Then the shortest path in $\mathcal{G}$ from $\bar{\mathbf{x}}$ to † has a total edge length equal to the duration until the time at which the control loop should next be closed. This shortest path can be computed using Dijkstra's algorithm, for example. For any graph $\mathcal{G}$ constructed using the above method, the shortest possible path will be strictly positive since the time spent in the original cell $\bar{\mathbf{x}}$ under the control $\bar{u}$, i.e., the interpolation interval $\widetilde{\Delta t}(\bar{\mathbf{x}}, \bar{u})$, is strictly positive.

In lieu of a proof, we now describe the reasoning behind this algorithm. Instead

of confirming stability by checking if $e^{-\beta t}V(\mathbf{x})$ is decreasing, we examine whether the interpolated parameter process $\tilde{\xi}_n^h$, which is locally-consistent with $e^{-\beta t}V(\mathbf{x})$, is decreasing in expected value under the outdated sample $\bar{\mathbf{x}}$ and control $\bar{u}$. The grid cells $\mathbf{x}_i$ where this relation is violated, i.e., where $\mathrm{d}V(\mathbf{x},t) \geq 0$, satisfy

$$-\beta V(\mathbf{x}_i,t) + f(\mathbf{x}_i,\bar{u})^T \nabla V(\mathbf{x}_i) + \frac{1}{2} g(\mathbf{x}_i)^T \nabla^2 V(\mathbf{x}_i,t) g(\mathbf{x}_i) \geq 0.$$

Adding zero in the form of $f(\mathbf{x}_i,u_i)^\intercal \nabla V(\mathbf{x}_i) - f(\mathbf{x}_i,u_i)^\intercal \nabla V(\mathbf{x}_i) + k(\mathbf{x}_i,u_i) - k(\mathbf{x}_i,u_i)$, we obtain

$$\begin{aligned} -\beta V(\mathbf{x}_i,t) + f(\mathbf{x}_i,u_i)^\intercal \nabla V(\mathbf{x}_i) + \frac{1}{2} g(\mathbf{x}_i)^T \nabla^2 V(\mathbf{x}_i,t) g(\mathbf{x}_i) + k(\mathbf{x}_i,u_i) \\ + \left(f(\mathbf{x}_i,\bar{u}) - f(\mathbf{x}_i,u_i)\right)^T \nabla V(\mathbf{x}_i) \geq k(\mathbf{x}_i,u_i). \end{aligned}$$

From the HJB equation, which was used to compute $V(\mathbf{x}_i)$ and u_i, the first four terms cancel, leaving the condition (3.21). The minimum amount of time before stability is lost (in the interpolated Markov chain) is given by the total amount of time spent in the chain until the condition (3.21) is first met. Since the edge lengths are equivalent to the interpolation intervals, the total edge length in the graph $\mathcal{G}$ is equivalent to the amount of time before stability of the chain $\xi^h(t)$ is lost[5]. The shortest path to † in the chain is only one possible path, of course, and with some probability the system could remain stable for a longer period of time; this algorithm is, therefore, conservative.

[5]Note that the chain constructed using the method of the previous sections has the property that $p(\mathbf{x}|\mathbf{x},u) = 0$. Consequently, the amount of time spent in the state $\mathbf{x}_i$ is exactly $\Delta t(\mathbf{x}_i,\bar{u})$, which is deterministic.

Chapter 4

A Stochastic Approach to Dubins Vehicle Tracking Problems

This chapter is a preprint to the paper

- Anderson, R. P., and Milutinović, D., "A Stochastic approach to Dubins vehicle tracking problems," *IEEE Transactions on Automatic Control*, (in press).

4.1 Introduction

In path planning and trajectory optimization problems, an autonomous robot may execute a task with limited or no knowledge of the future effects of its immediate actions [246]. For example, an Unmanned Aerial Vehicle (UAV) may wish to track, protect, or provide surveillance of a ground-based target. If the target trajectory is known, a deterministic optimization or control problem can be solved to give a feasible UAV trajectory. Our goal in this work is to develop a feedback control policy that allows a UAV to optimally maintain a nominal standoff distance from the target *without full knowledge of the current target position or its future trajectory*.

The UAV is assumed to fly at constant altitude and with a bounded turning rate. In

this work, its kinematics is described by that of a planar Dubins vehicle [82], which gives a good approximation for small fixed-wing UAV trajectories. The Lyapunov stability-based control design [197] is commonly used to develop feedback controllers for problems of this type [139, 211, 288], but constructing a Lyapunov function may not be a straightforward task. Alternatively, the control problem may be defined as an infinite-horizon optimal control problem, resulting in a Hamilton-Jacobi-Bellman (HJB) partial differential equation (PDE), whose solution, the value of the cost-to-go function, serves as a Lyapunov function that can be constructed computationally.

In order to account for unknown target kinematics, we assume that the target motion can be robustly described by planar Brownian motion [274]. Although, strictly speaking, our tracking control is optimal in the expected value sense for the random walk, it can be applied to a wider class of continuous and smooth target trajectories. Moreover, since observations of the target may be disrupted in a realistic scenario due to communication channel constraints, unexpected terrain, signal corruption, or asymmetric observation capabilities, for example, we allow for a tracker to probabilistically “lose" its target and later locate it. In our approach, the amount of time since the last target observation is a state variable, and the change in position of the target upon observation gives rise to a stochastic jump process [183]. The possibility of finding the target in a costly location adds an implicit motivation for the tracker to optimally position itself in anticipation of future observations.

Stochastic problems in the control of Dubins vehicles typically concentrate on vari-

ants of the Traveling Salesperson Problem and other routing problems, in which the target location is unknown or randomly-generated [90, 223, 225]. One exception is [201] in which the authors apply a sample-based dynamic programming approach to UAV flocking, but where the target (in this case, the leader UAV) has known control and is continuously observed. In this article, our focus is on a novel stochastic formulation of the problem when this information is unavailable, and on the corresponding numerical feedback control solution[1]. Our feedback control policy is computed off-line using a Bellman equation discretized through an approximating Markov chain [137]. This method is well-accepted and necessary for an accurate discretization, but it is rarely seen in the robotics community.

In what follows, we first describe our stochastic modeling approach in Section 4.2. In Section 4.3 we describe how to compute the optimal feedback control and analyze the connection between the parameters of the problem and the resulting control law. In Section 4.4 we demonstrate the effectiveness of this approach for both Brownian targets, targets with unknown deterministic trajectories, and for the case where observations can be lost. Section 4.5 concludes this paper and provides directions for future research.

[1]Preliminary versions of this work have appeared in [10, 11, 15].

4.2 Problem Formulation

We consider a small, fixed-wing UAV flying at a constant altitude in the vicinity of a ground-based target, tasked with maintaining a nominal distance from the target. The target is located at position $[x_T(t),\, y_T(t)]^T$ at the time point t (see Fig. 4.1). The UAV, located at position $[x_A(t),\, y_A(t)]^T$, moves in the direction of its heading angle θ at a constant speed v.

The UAV model is that of a planar Dubins vehicle [82]:

$$\begin{aligned} \mathrm{d}x_A(t) &= v\cos(\theta(t))\,\mathrm{d}t \\ \mathrm{d}y_A(t) &= v\sin(\theta(t))\,\mathrm{d}t \\ \mathrm{d}\theta(t) &= u\mathrm{d}t, \qquad u \in \mathcal{U} \end{aligned} \tag{4.1}$$

where the turning rate $u \in \mathcal{U} \equiv \{u : |u| \leq u_{\max}\}$ must be found. This is a common model for UAVs with autopilots sufficiently capable of achieving a commanded heading angle $\theta_c = \theta + u/\alpha$ [198, 205], so that the evolution of θ could be alternatively described by a first-order differential equation $\mathrm{d}\theta(t) = \alpha(\theta_c - \theta)\,\mathrm{d}t$, where $\alpha > 0$ is a system parameter.

In our problem formulation, the target motion is unknown. We therefore assume that it is random and described by a 2D stochastic process. Drawing from the field of estimation, the simplest signal that can be used to describe an unknown trajectory

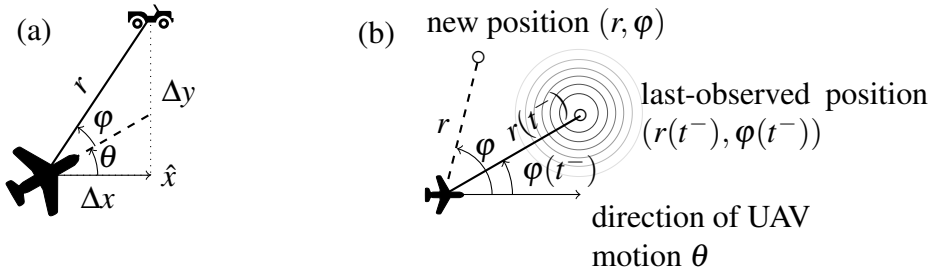


Figure 4.1: (a) Diagram of a UAV that is moving at heading angle θ and tracking a randomly-moving target with distance r and relative angle φ. (b) Contour representation of probability density function for observing the target in a new position (r,φ) after a previous observation at $(r(t^-),\varphi(t^-))$.

suggests that the motion of the target should be described by 2D Brownian motion:

$$\mathrm{d}x_T(t) = \sigma \mathrm{d}w_x, \qquad \mathrm{d}y_T(t) = \sigma \mathrm{d}w_y \tag{4.2}$$

where $\mathrm{d}w_x$ and $\mathrm{d}w_y$ are increments of unit intensity Wiener processes along the x and y axes, respectively, which are mutually independent. The level of noise intensity σ determining the target motion is assumed to be in a range that allows the UAV to effectively pursue or maintain pace with the target with high probability (this can be checked using, for example, the properties of a Rayleigh distribution [51]). For simplicity, any additional noise due to UAV observation error is also incorporated into the random target motion, i.e., the parameter σ, which is assumed to be constant over the observation area.

We frame the problem in terms of relative dynamics based on a time-varying coordinate system aligned with the direction of the UAV velocity. Defining $\Delta x = x_T - x_A$ and $\Delta y = y_T - y_A$, the reduced system state is composed of the distance between the UAV and target $r = \sqrt{(\Delta x)^2 + (\Delta y)^2}$ and the viewing angle $\varphi = \tan^{-1}(\Delta y/\Delta x) - \theta$, as

seen in Fig. 4.1(a).

We aim to construct a feedback control policy $u(r,\varphi)$ based on the UAV's knowledge of the target position in (r,φ), but this position may in reality not be observed continuously. Moreover, in many circumstances, the length of time between observations is not known in advance. Consequently, we address here the control strategy both for continuous observations, and that in the presence of observation loss. Introducing the state variable $\tau(t)$ as the time since the last observation of the target, the state variables r and φ can then be interpreted as the distance and viewing angle, respectively, to the last-observed target position $(x_T(t-\tau), y_T(t-\tau))$. Since we are not dealing with the problem of initially locating the target, we assume that at $t=0$, the tracking vehicle is observing the target. In the case of continuous observations, τ is fixed at 0.

The tracking vehicle should achieve and maintain a nominal distance d to the target. To this end, we seek to minimize the expectation of an infinite-horizon cost functional $V(\cdot)$ with a discounting factor $\beta > 0$ and with penalty $\varepsilon \geq 0$ for control:

$$V(r,\varphi,\tau) = \min_{u\in\mathcal{U}} \mathbb{E}\left\{ \int_0^\infty e^{-\beta t}\, k(r(t),u)\, \mathbb{1}_{\{\tau=0\}}\, \mathrm{d}t \right\} \tag{4.3}$$

where $k(r,u) = (r-d)^2 + \varepsilon u^2$. The indicator function $\mathbb{1}_{\{\tau=0\}}$ ensures that cost is only accumulated when the UAV is currently observing the target, i.e., $\tau = 0$. If $\tau > 0$, the tracking vehicle may wish to position itself near the nominal distance in anticipation of an observation.

In the case of continuous observations (with fixed $\tau = 0$), the differentials dr and

$\mathrm{d}\varphi$ can be found using Itô's Lemma with (4.1)-(4.2) as

$$\mathrm{d}r = \left(-v\cos\varphi + \frac{\sigma^2}{2r}\right)\mathrm{d}t + \sigma\mathrm{d}w_0 \tag{4.4}$$

$$\mathrm{d}\varphi = \left(\frac{v}{r}\sin\varphi - u\right)\mathrm{d}t + \frac{\sigma}{r}\mathrm{d}w_\perp \tag{4.5}$$

$$\mathrm{d}\tau = 0 \tag{4.6}$$

where $\mathrm{d}w_0 = \cos(\theta+\varphi)\,\mathrm{d}w_x + \sin(\theta+\varphi)\,\mathrm{d}w_y$ and $\mathrm{d}w_\perp = -\sin(\theta+\varphi)\,\mathrm{d}w_x$ $+\cos(\theta+\varphi)\,\mathrm{d}w_y$ are independent Wiener process increments due to the invariance of noise under a rotation of the coordinate frame [105]. Note the appearance of a positive bias $\sigma^2/2r$ in the relation for $r(t)$, which is a consequence of the random process included in our analysis.

Turning to the possibility for observation loss, we define a rate λ_{12} of losing the target, so that, for small $\mathrm{d}t$, $\Pr(\text{target lost in } [t, t+\mathrm{d}t) \mid \text{observed}) = \lambda_{12}\mathrm{d}t + o(\mathrm{d}t)$. Then between observations, the last-observed target position does not change, and r and φ are described by the equations (4.4)-(4.6) with $\sigma = 0$. At the time of the observation, however, the state variables may change significantly from their previous values due to the random target motion. We model this behavior with a jump process $J(t) = \left[J_r(t), J_\varphi(t), J_\tau(t)\right]^T$. The kinematics model that accounts for observation loss is

$$\mathrm{d}r(t) = (-v\cos\varphi)\,\mathrm{d}t + \mathrm{d}J_r \tag{4.7}$$

$$\mathrm{d}\varphi(t) = \left(\frac{v}{r}\sin\varphi - u\right)\mathrm{d}t + \mathrm{d}J_\varphi \tag{4.8}$$

$$\mathrm{d}\tau(t) = \mathrm{d}t + \mathrm{d}J_\tau. \tag{4.9}$$

A formal definition of the jump process is possible, but for the purpose of describing our

approach, we only require the probability that a jump occurs in any small time interval and the probability of a new state after a jump. We define a rate λ_{21} of finding the target after observation failure, so that, for small $\mathrm{d}t$, $\Pr(\text{target found in } [t, t+\mathrm{d}t) \mid \text{lost}) = \lambda_{21}\mathrm{d}t + o(\mathrm{d}t)$. If a jump due to a target observation occurs in $[t, t+\mathrm{d}t)$, the jump from $[r(t^-), \varphi(t^-), \tau(t^-)]^T$ to $[r(t), \varphi(t), 0]^T$ then encodes the cumulative random motion between the target and the vehicle in state space during a loss of observation of duration τ, and it also resets τ to $\tau = 0$. The joint distribution of the components $\mathrm{d}J_r$, $\mathrm{d}J_\varphi$, and $\mathrm{d}J_\tau$ is based on the probability of the resulting state, which we denote by $p(r, \varphi | r(t^-), \varphi(t^-), \tau(t^-)) \times \delta(\tau)$ and derive in the Appendix A. Comparing (4.7)-(4.9) to (4.4)-(4.6), note that both the bias term $\sigma^2/2r$ and the effects of the Wiener process have been incorporated into $J(t)$. In summary, without observation, Equations (4.7)-(4.9) describe the motion of the observed state, which is the "relative position to the last target observation." Equations (4.4)-(4.6) describe the motion of the same state but under continuous observation, that is, the relative position to the last target observation is continuously updated.

To determine the set of admissible control $\mathcal{U}$, we use (4.4)-(4.6) in the limit of no noise ($\sigma = 0$). With the steady state of these equations in mind, the admissible set is given by $|u| \leq v/r_{\min} \equiv u_{\max}$ for the nominal distance $d \in (r_{\min}, r_{\max})$.

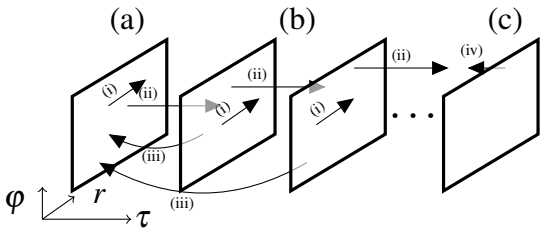


Figure 4.2: Allowed transition probabilities based on the time since observation τ. Illustrated panels are (a) $\tau = 0$, (b) $0 < \tau < \tau_{\max}$, and (c) $\tau_{\max}$. Arrows indicate (i) spatial transition probabilities (ii) increments in variable τ, (iii) jumps due to target observation, and (iv) reflection at $\tau = \tau_{\max}$.

4.3 Optimal Controllers

4.3.1 Computing the Control

For an accurate state space discretization for dynamic programming, we employ the *Markov chain approximation method* [137] for numerically determining the optimal control policy corresponding to the controlled diffusion processes (4.4)-(4.6) and (4.7)-(4.9) and cost function (4.3). We now describe how one can apply these methods for the problem at hand.

The state space of the control problem is represented in Fig. 4.2. The allowed transitions among states vary based on the value of τ. In the case of continuous observations (with fixed $\tau = 0$), transitions occur only in r and φ, and the dynamic programming equation for value iterations can be found using (4.4)-(4.5) and [137, pp. 106–113] for states inside the computational domain. The domain boundary $r = r_{\min}$ and $r = r_{\max}$

is reflective, and for the states in this boundary, we use [137, p. 143][2]. The domain is periodic in φ.

For the case of exponentially-distributed observation durations, we extend the ideas of [137, Ch. 12] and devise a value iteration algorithm that includes the time variable τ. The equation for value iterations when $\tau = 0$ includes the state space transitions due to (4.4)-(4.6), but with rate λ_{12}, the state will increase in τ by $\Delta\tau$, i.e., with an explicit treatment of τ. For $0 < \tau < \tau_{\max}$, since the time-like variable τ is reset to 0 upon an observation, we must also ensure consistency with respect to the jump process (4.7)-(4.9). Based on λ_{21}, the state will either jump to the $\tau = 0$ plane with a new state given by $p(r, \varphi | r(t^-), \varphi(t^-), \tau(t^-))$, in which case the equations for value iterations can be found using [137, pp. 127–132], or it will not receive an observation, in which case transition probabilities in both (r, φ) and τ may be found using an implicit approximation method [137, pp. 333–337]. Finally, we choose the state $\tau = \tau_{\max}$ to be reflective, since, for large $\tau_{\max}$, the jump distribution becomes practically spatially uniform, and the probability of reaching $\tau = \tau_{\max}$ is slim.

Based on these transition probabilities, we use the standard method of value iteration until the cost converges. This required approximately 10 seconds for fixed $\tau = 0$ and 48 hours for the case of observation loss on an Intel i7 with 8GB of RAM. In the examples, the control is obtained by interpolating the current system state to the discretized control $u(r, \varphi, \tau)$.

[2]The dynamic programming update equation for these states is similar to that found in Appendix B, Equation (B.8) but with γ replaced by $\tau = 0$. (Note: this footnote does not appear in the original manuscript)

4.3.2 Control under continuous observations

Here we describe the control computed by dynamic programming for various system parameters. In all cases, the nominal distance is 50, $r_{\min} = 10$ and $r_{\max} = 90$, and the UAV velocity v is 10, with all units in meters and seconds. The target noise intensity is $\sigma = 5$, and discount factor $\beta = 1$.

For the cost function (4.3), if $\varepsilon = 0$, there is no penalty for control, and the optimal turning rate for the Dubins vehicle as computed by the value iterations method is given by a bang-bang[3] controller $u(r,\varphi) \in \{-u_{\max}, u_{\max}\}$, as seen by the sharp edges in Fig. 4.3 for $\sigma = 5$. As such, it is highly responsive to the random motion of the target, and based on previous works, this type of controller is not unexpected [82], [243]. Notably, due to the additive noise, the state will almost never lie on the switching curve, and, therefore, the control structure is well defined. If $\varepsilon = 1$, the penalty smoothes the transitions among the regions. Open regions away from these boundaries have the effect of directing the state back to the lines. If the vehicle is in state $(80\text{ m}, -\pi/30)$, for example, it is far from the target and directed toward it, but with a small angular offset that hints at the future rotation about its target. As it approaches the target, the curvature of the region gradually directs the vehicle into a circle about the target, beginning at $r = r_2$. This continues until the vehicle reaches the steady pattern. The positioning of r_1 and r_2 determine the state at which the Dubins vehicle transitions

[3]The control is that which switches between one extreme $u_{\max}$ and the other $-u_{\max}$ (Note: this footnote does not appear in the original manuscript)

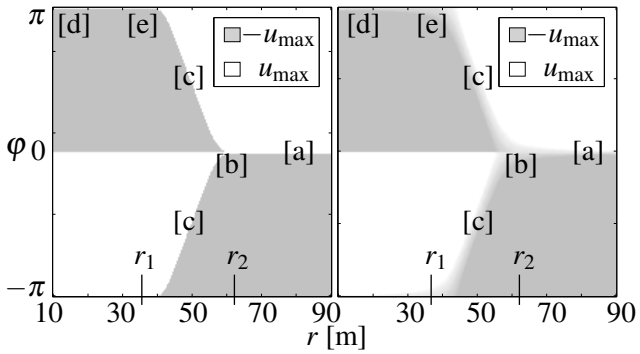


Figure 4.3: Optimal control based on distance to the target r and viewing angle φ for $\sigma = 5$, $\varepsilon = 0$ (left) and $\varepsilon = 1$ (right). Indicated points are [a] heading toward target at a small angle displacement from a direct line, [b] start of clockwise rotation about target, [c] steady states at $\left(d, \pm\cos^{-1}\left(\sigma^2/100v\right)\right)$, [d] heading directly away from target at a small angle displacement away from a direct line, [e] start of counter-clockwise rotation about target.

from the act of “avoiding" or “chasing" the target when it is too close or too far, respectively, into the act of circling the target. Owing to the bias in the mean drift of dr (4.4), $|r_1 - d| > |r_2 - d|$, i.e., a Dubins vehicle avoiding the target must anticipate the bias to “help" with this action, while when chasing, the bias might hinder these efforts. The effect of σ on this phenomenon is seen in Fig. 4.4.

In simulations, the actual evolution of the state (r, φ) is not smooth due to the random motion of the target. Since we began this problem with an infinite-horizon cost, the control is highly robust to such deviations, as we will show in simulations.

4.3.3 Control with observation interruptions

The control computed for the case of observation loss may be found in Fig. 4.5 for $\lambda_{12} = 0.01$ / $\lambda_{21} = 0.1$ and $\tau_{\max} = 60$ s, and in Fig. 4.6 for $\lambda_{12} = 0.05$ / $\lambda_{21} = 0.5$. When $\tau = 0$ s, the control policy matches that for tracking a stationary target. However,

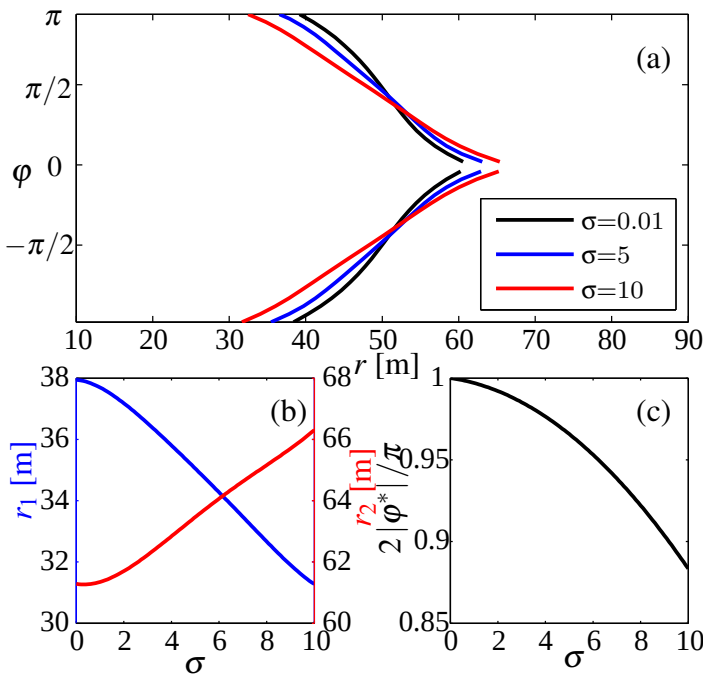


Figure 4.4: (a) Switching curves from Fig. 4.3 for various values of σ. (b) Radii r_1 and r_2 at which the UAV begins its entrance into a turning pattern about the target, corresponding to the labeled regions in Fig. 4.3 for $\sigma = 5$. As the noise intensity of the target increases, the UAV has less information of where the turning circle should be centered and must begin its turn sooner. (c) Coordinate $\varphi^\star$ of switching boundary intersection with $r = d$. As the target noise increases, the UAV expects a bias that tends to increase the target distance, and it reduces its steady-state viewing angle accordingly.

for $\tau > 0$, the control instructs the tracker to gradually spiral in toward the position where the target was last spotted until it reaches its minimum turning radius, as illustrated (simulated) in Fig. 4.7. This change is likely due to the fact that the angular component of any radial displacement of a symmetric Brownian motion will be uniformly distributed. Once the target sufficiently deviates from its original position, it becomes advantageous for the Dubins vehicle to center itself nominally to all the target's possible positions. The spiral curvature is smaller for larger λ_{12}.

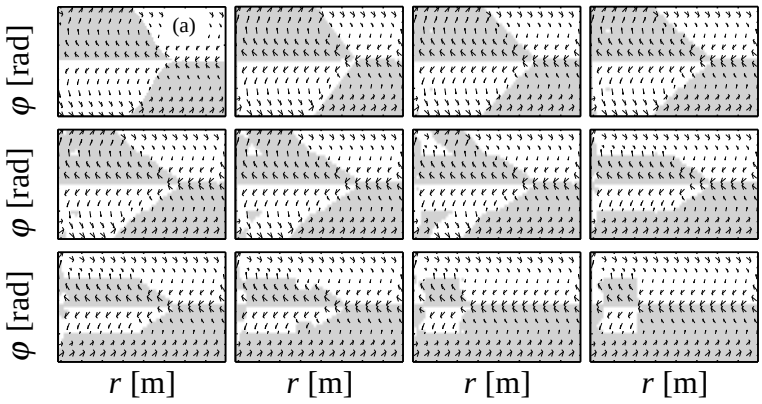


Figure 4.5: Control anticipating observation loss, $\lambda_{12} = 0.01$ / $\lambda_{21} = 0.1$. Control assuming continuous observations is listed for comparison in (a). From left-to-right, top-to-bottom, $\tau = 0$, 2.1, 8.6, 15.0, 21.4, 27.9, 34.3, 40.7, 47.1, 53.6, and 57.9 s, respectively. Color mapping is the same as in Fig. 4.3.

4.4 Simulations

We show implementations of this control when the UAV is initially too far or too close to the target in Fig. 4.8(a-d). Since a small time-step was used for simulation, the Dubins vehicle trajectory using the control penalty $\varepsilon = 0$ was indistinguishable from when $\varepsilon = 1$, but since its response was less sensitive to small displacements in target location when $\varepsilon = 1$, the average standoff distance was slightly affected.

During observation loss, the control computed anticipating this provides a mean distance that is nearly equal to the nominal distance (49.87 m for $\lambda_{12} = 0.01$ / $\lambda_{21} = 0.1$ and 49.83 m for $\lambda_{12} = 0.05$ / $\lambda_{21} = 0.5$), while the control assuming continuous observations tends to be too far to the target during this time (53.21 m for $\lambda_{12} = 0.01$ / $\lambda_{21} = 0.1$ and 51.68 m for $\lambda_{12} = 0.05$ / $\lambda_{21} = 0.5$). However, the performance decrease is not large, indicating that, for this problem, the use of an outdated target observation

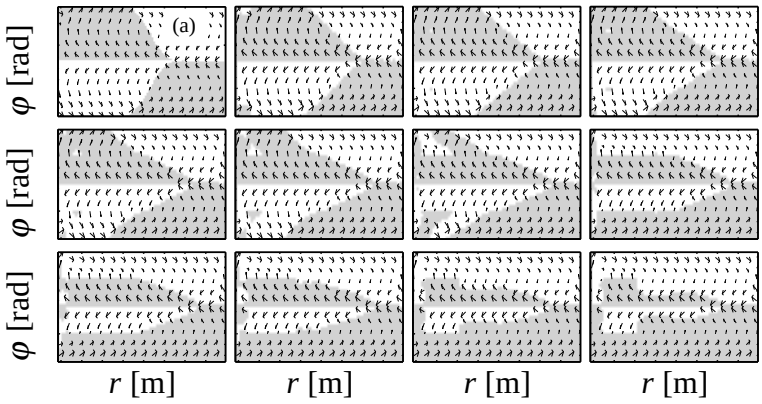


Figure 4.6: Control anticipating observation loss, $\lambda_{12} = 0.05$ / $\lambda_{21} = 0.5$. From left-to-right, top-to-bottom, the values of τ match those in Fig. 4.5.

may suffice.

Next, to emphasize the level of robustness provided by the control, the original assumption that the target position evolves as a 2D random walk is dropped. We exhibit the response to a target moving in a complex sinusoidal path with speed 4 m/s in Fig. 4.9. Since target motion is no longer random, our control is no longer strictly optimal. It is seen in Fig. 4.9(b), however, that the vehicle remains near the nominal standoff distance, with the average distance slightly below d since the bias in $r(t)$ is no longer present. Finally, we exhibit the response to the same trajectory under observation loss ($\lambda_{12} = 0.01$ / $\lambda_{21} = 0.1$).

4.5 Conclusion and Future Work

This paper considers the problem of maintaining a nominal distance between a Dubins vehicle / UAV and a ground-based target. Brownian motion serves as a prior for the

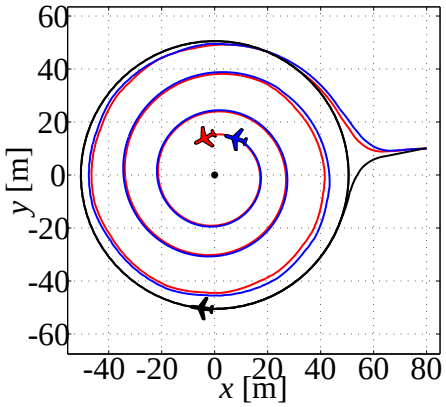


Figure 4.7: Comparison of control policy under observation loss to continuous observation control policy. The target, fixed at $(0,0)$, is only observed at $t = 0$. The black Dubins vehicle applies the control for a continuously-observed random walk target, while the red Dubins vehicle applies the control anticiating observation loss with $\lambda_{12} = 0.01$ / $\lambda_{21} = 0.1$. The blue Dubins vehicle applies the same control but for $\lambda_{12} = 0.05$ / $\lambda_{21} = 0.5$. The direction of rotation is a consequence of the initial condition and the symmetry breaking that occurs during value iterations to pick a single direction, i.e., clockwise or counter-clockwise.

target kinematics, and the possibility for a loss of observation is included using probabilistic jumps. A Markov chain approximation that is locally consistent with the system under control is constructed on a discrete state space, and value iterations on the associated cost function produce a UAV turning rate control to minimize the expected mean squared distance to the target in excess of a nominal distance.

The off-line control needs to only be computed once for a given target noise, regardless of trajectory shape or initial conditions, and takes into account kinematic nonlinearities. The assumption of random target dynamics coupled with an infinite horizon cost has the fundamental advantage of creating a highly robust control. A variety of trajectories can be tracked using this approach, despite the fact that the intensity of the noise in the UAV-target distance is no longer present in the case of natural motion.

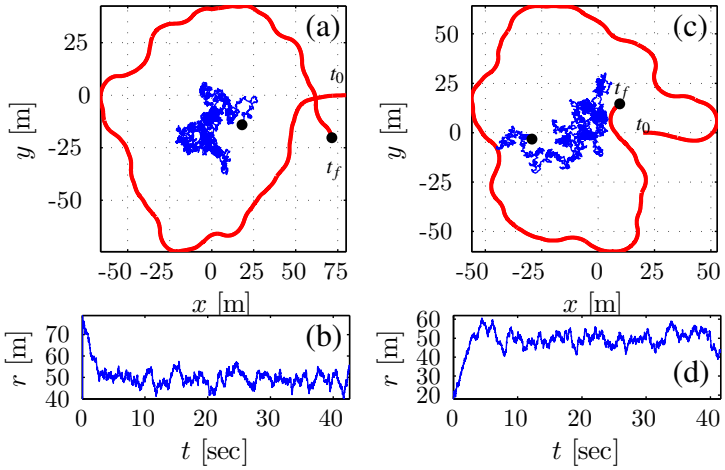


Figure 4.8: A Dubins vehicle (red) tracking a Brownian target (blue). (a) For $\varepsilon = 0$, the Dubins vehicle must approach the target ($t_0 = 0$) before entering into a clockwise circular pattern ($t_f = 42.6$) and (b) its associated distance r ($mean(r) = 50.23$, $std(r) = 4.97$). For $\varepsilon = 1$, $mean(r) = 49.03$, $std(r) = 5.02$ (c) The Dubins vehicle begins near the target ($t_0 = 0$) and must first avoid the target using $\varepsilon = 0$ before beginning to circle ($t_f = 41.6$). The position of the target jumps sharply to the East near the end of the simulation, and the Dubins vehicle is shown turning right to avoid it. (d) The associated distance ($mean(r) = 48.8$, $std(r) = 5.98$). For $\varepsilon = 1$, $mean(r) = 48.7$, $std(r) = 5.94$.

When approaching the nominal standoff distance from too close or too far a proximity, the UAV flight pattern that initializes revolutions about the target and the revolutions themselves are determined by the anticipated variability of the target, even if its motion is smooth and deterministic. The computed control policies indicate that if some time has passed since the last target observation, the tracker should spiral in toward the last-observed target position. These conclusions are valid only for the parameter regimes studied herein, and it is not known at this time if other behaviors could arise in different regimes. Having faced some numerical instabilities for high observation loss rates, the study of this question may require alternative numerical methods, as higher observation loss rates may cause the problem to verge on open loop control territory.

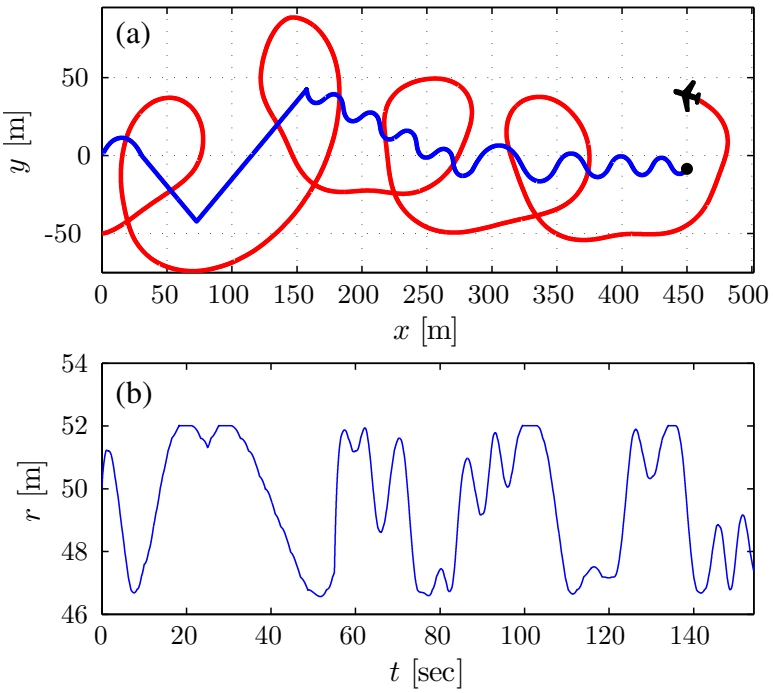


Figure 4.9: A UAV (red) tracking a target (blue) moving in a complex sinusoidal path using the control for a Brownian target. (a) The UAV follows the target in eccentric circles whose shape is determined by the current position of the target along its trajectory, $t \in [0, 154.2]$. (b) The distance r ($mean(r) = 49.5$ m, $std(r) = 3.74$).

Should the target also be modeled as a Dubins vehicle with a Brownian heading angle, the knowledge of the target's heading angle would provide the UAV with an indicator of the target's immediate motion and an appropriate response (control). Extensions to state-dependent or correlated noise[4] are also possible, although this would complicate the numerical method and potentially increase the state space dimension.

[4]Filtered noise would increase the dimension of the state space (Note: this footnote does not appear in the original manuscript)

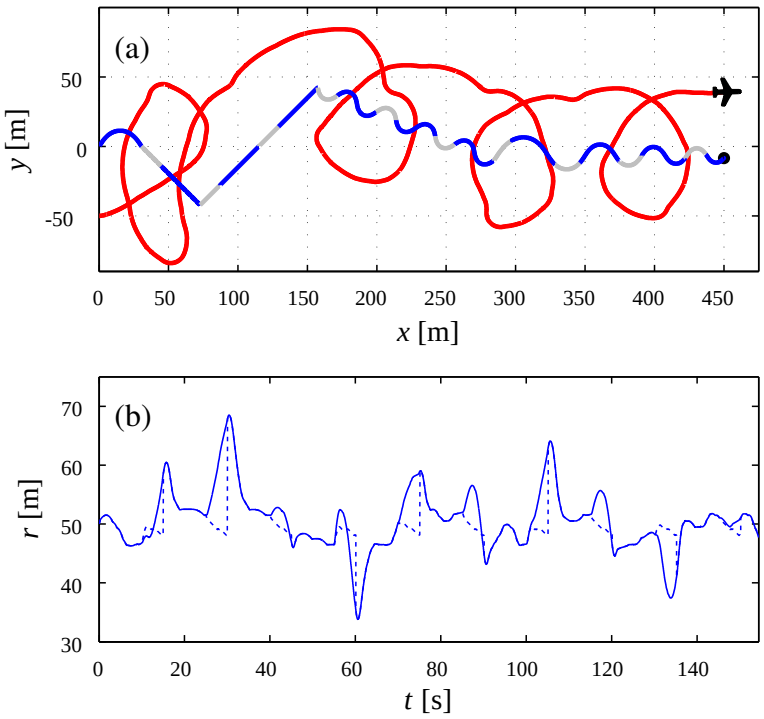


Figure 4.10: (a) A UAV (red) tracking a target using the control for $\lambda_{12} = 0.01$ and $\lambda_{21} = 0.1$ when alternating between observing the target (blue) for 10 s and not observing the target (gray) for 5 s. (b) The associated distances are the true distance (solid; $mean(r) = 50.6$ m, $std(r) = 50.6$) and the distance to the last-observed target position (dashed; $mean(r) = 49.8$ m, $std(r) = 3.9$).

Chapter 5

Optimal Feedback Guidance of a Small Aerial Vehicle in the Presence of Stochastic Wind

This chapter is a preprint to the paper

- Anderson, R. P., Bakolas, E., Milutinović, D., and Tsiotras, P., "Optimal Feedback Guidance of a Small Aerial Vehicle in a Stochastic Wind", *AIAA Journal of Guidance, Control, and Dynamics*, vol. 36, issue 4, pp. 975-985, 2013.

5.1 Introduction

This paper deals with the problem of guiding an aerial vehicle with a turning rate constraint to a prescribed terminal position in the presence of a stochastic wind in minimum expected time. It is assumed that the motion of the vehicle is described by a Dubins-like kinematic model [82, 118, 191], that is, it travels only forward with constant speed, and such that the rate of change of its forward velocity vector direction is bounded by a prescribed upper bound. This kinematic model is referred to as the Dubins Vehicle (DV for short). In the absence of the wind, the vehicle traverses paths of minimal length

and bounded curvature, known in the literature as Dubins paths or optimal paths of the Markov-Dubins (MD for short) problem [82, 244].

The importance of designing trajectory tracking control schemes and path planning algorithms that account for the effects of the local wind in UAVs/MAVs applications has been recognized by many researchers. In particular, Refs. [61, 178, 215, 295] present path tracking/following controllers for UAVs/MAVs in the presence of disturbances induced by the wind based on nonlinear control tools. The problem of characterizing minimum-time paths of a DV in the presence of a constant wind was first posed by McGee and Hedrick in [166]. Numerical schemes for the computation of the Dubins-like paths proposed in [166] have been presented in [26, 253]. The complete characterization of the optimal solution of the same problem, that is, a mapping that returns the minimum-time control input given the state vector of the DV, is given in [24, 26]. A numerical algorithm that computes the minimum-time paths of the DV in the presence of a deterministic time-varying, yet spatially invariant, wind is presented in [168].

The analysis presented in the majority of the variations and extensions of the MD problem in the literature is based on a deterministic optimal control framework (the reader interested in a thorough literature review on variations/extensions of the MD problem may refer to [28] and references therein). The effect of the wind, however, is intrinsically stochastic, and approaching this problem from a stochastic point of view is more appropriate. Some recent attempts to address optimal control problems related to the MD problem within a stochastic optimal control framework can be found in [10, 11].

In particular, Refs. [10, 11] deal with the problem of a DV tracking a target with an uncertain future trajectory using numerical techniques from stochastic optimal control of continuous-time processes [137].

In this work, an optimal *feedback* control that minimizes the expected time required to navigate the DV to its prescribed target set in the presence of a stochastic wind is developed. Two stochastic wind models are investigated. In the first model the x and y components of the wind are modeled as independent zero-mean Wiener processes with a given intensity level. In the second model, the wind is modeled as having a constant magnitude, but its direction is unknown and is allowed to vary stochastically according to a zero-mean Wiener process. For both wind models, optimal feedback control laws are computed numerically using a Markov chain approximation scheme. In addition, for each control based on the stochastic wind models, feedback control laws based on deterministic wind models are developed and compared against their stochastic model-based counterparts in the presence of stochasticity to determine the regions of validity of the former. The analysis and numerical simulations demonstrate, not surprisingly perhaps, that control laws based on stochastic wind models outperform – on the average – control laws for the deterministic wind models implemented in a stochastic wind. On the other hand, the control laws for the deterministic wind model can successfully capture the salient features of the structure of the corresponding stochastic optimal control solution.

The contributions of the paper can be summarized as follows: First, this paper

offers, up to the authors' best knowledge, for the first time, the solution of the optimal path generation of an aerial vehicle with Dubins-like kinematics in the presence of stochastic wind. This is important for small UAV path-planning and coordination applications. Second, it shows the relationship of the optimal control solution, which anticipates the stochastic wind, with its deterministic counterpart, and it compares the two. This allows one to draw insights as to what level a stochastic wind-based solution is beneficial compared to its less informed deterministic counterpart, and when it makes sense (from a practical point of view) to use the former over the latter. The question of the use of a feedback control law anticipating stochastic processes versus a control law based on deterministic model assumptions is a question of a more general interest and one that is a recurrent theme in the community, especially in terms of applications. This paper offers a rare example where a head-to-head comparison is possible. In general, the computation of a deterministic optimal feedback control is not an easy task, as it requires the solution of a Hamilton-Jacobi-Bellman (HJB) partial differential equation. However, the Dubins vehicle problem in this paper serendipitously allows for a complete solution, via a synthesis of open-loop strategies, without resorting to the HJB equation.

The rest of the paper is organized as follows. Section 5.2 formulates the optimal control problem. Section 5.3 presents feedback control laws based on minimal deterministic and stochastic wind model assumptions. These control laws are extended in Section 5.4 for the case when the wind has a known speed but stochastically-varying

direction. Simulation results for the controllers based on the two types of deterministic and stochastic wind models are presented in Section 5.5. Finally, Section 5.6 concludes the paper with a summary of remarks.

5.2 Problem Formulation

Here the problem of controlling the turning rate of a fixed-speed Dubins vehicle (DV) in order to reach a stationary target in the presence of wind is formulated. The target is fixed at the origin, while the Cartesian components of DV position are $x(t)$ and $y(t)$ (see Fig. 5.1).

The DV moves in the direction of its heading angle θ at fixed speed v relative to the wind and obeys the equations:

$$\mathrm{d}x(t) = v\cos(\theta)\,\mathrm{d}t + \mathrm{d}w_x(t,x,y), \tag{5.1}$$

$$\mathrm{d}y(t) = v\sin(\theta)\,\mathrm{d}t + \mathrm{d}w_y(t,x,y), \tag{5.2}$$

$$\mathrm{d}\theta(t) = \frac{u}{\rho_{\min}}\mathrm{d}t, \qquad |u| \leq 1, \tag{5.3}$$

where $\rho_{\min} > 0$ is the minimum turning radius constraint (in the absence of wind) and u is the control variable, $u \in [-1,1]$. The motion of the DV is affected by the spatially and/or temporally varying wind $w(t,x,y) = [w_x(t,x,y), w_y(t,x,y)]^{\mathsf{T}}$, whose increments have been incorporated into the model (5.1)-(5.2). In this problem formulation, the model for the wind is unknown. Therefore, it is assumed that the wind is described by a stochastic process. Subsequently, a stochastic control problem for reaching a target

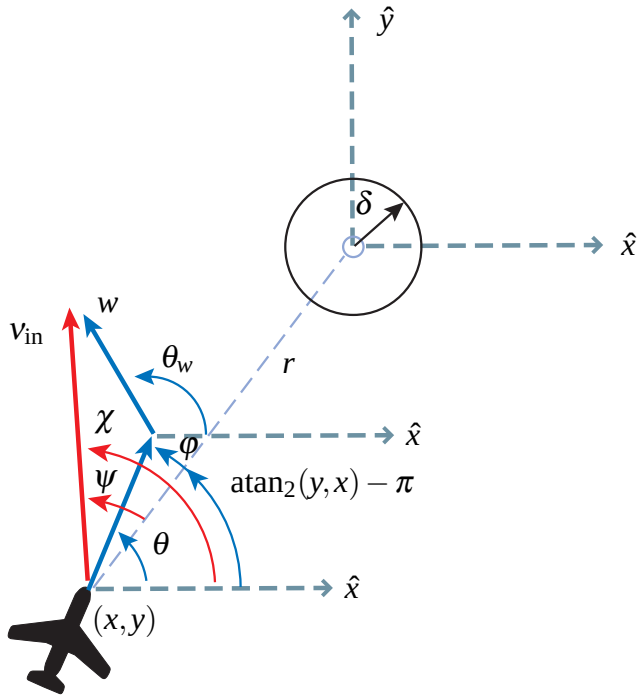


Figure 5.1: Diagram of a DV at position $[x(t),\, y(t)]^\mathsf{T}$ moving at heading angle θ in order to converge on a target in minimum time in the presence of wind. The target set $\mathcal{T}$ is shown as a circle of radius δ around the target. In the presence of the wind vector w at an angle θ_w, the DV travels in the direction of its inertial velocity vector v_{in} and with a line-of-sight angle φ to the target center. The angle between v_{in} and the line-of-sight angle is φ, and χ is the angle $\text{atan}_2(\dot{y},\dot{x})$.

set $\mathcal{T} = \left\{(x,y) \,:\, x^2 + y^2 \leq \delta^2\right\}$, which is a ball of radius $\delta > 0$ around the target, is formulated.

In order to minimize the time required to reach the target set, one defines a cost-to-go function

$$J(\mathbf{x}) = \min_{|u|\leq 1} \mathbf{E}\left[\int_0^T \mathrm{d}t\right], \qquad \mathbf{x} := [x(0), y(0), \theta(0)]^\mathsf{T}, \tag{5.4}$$

and assumes that upon reaching the target set $\mathcal{T}$ at time T, all motion ceases. In (5.4) the expected time to reach the target set is minimized over the turning rate u, $|u| \leq 1$, which

is the control variable. Control problems with a cost-to-go function of the form (5.4) are sometimes referred to as "control until a target set is reached" [137] or stochastic shortest-path problems [46].

Two stochastic process wind models that are characterized by the amount of information known about the wind are considered. In each case, it is assumed that the wind is a continuous-time stochastic process with respect to the DV position. In other words, there is no explicit relation between a realization of the wind and the DV position, i.e., $w_x(t,x,y) = w_x(t)$, and $w_y(t,x,y) = w_y(t)$, although implicitly this relation may exist.

First, a feedback control when a model describing the wind is not given is developed. Drawing from the field of estimation, the simplest model to describe an unknown 2D signal suggests that the wind should be modeled as Brownian motion [105]. It is further assumed that the Cartesian components of the wind evolve independently. Then from (5.1)-(5.3), the kinematics of the DV in the presence of this wind, denoted model (W1), is

$$\begin{aligned} \mathrm{d}x(t) &= v\cos(\theta)\mathrm{d}t + \sigma_W \mathrm{d}W_x, \\ \mathrm{d}y(t) &= v\sin(\theta)\mathrm{d}t + \sigma_W \mathrm{d}W_y, \\ \mathrm{d}\theta(t) &= \frac{u}{\rho_{\min}}\mathrm{d}t, \qquad |u| \leq 1, \end{aligned} \tag{W1}$$

where $\mathrm{d}W_x$ and $\mathrm{d}W_y$ are mutually independent increments of a zero mean Wiener process, and where the level of noise intensity σ_W quantifies the uncertainty in the evolution of the wind. Note that this kinematic model also arises when examining the problem of tracking a target with unknown future trajectory [10, 11]. In the limiting

case where $\sigma_W = 0$, the problem is reduced to the case without the wind. A feedback control that assumes $\sigma_W = 0$, therefore, would ignore the presence of the wind, while a feedback control that assumes $\sigma_W > 0$ will account for the stochastic wind variation. Along these lines, Section 5.3 develops feedback control laws that drive the DV to the target in minimum time in both the deterministic case ($\sigma_W = 0$) and the stochastic case ($\sigma_W > 0$). Note that in the deterministic case, the cost-to-go function is the same as (5.4), but without the expectation operator.

Next, motivated by problems involving a wind that varies slowly in time and/or space, a second wind model (W2) is considered. The second wind model considers a wind that flows in the direction θ_w at constant speed $v_w < v$, but where the evolution of the direction of this wind is unknown. Then from (5.1)-(5.3), the model of the relative motion of the DV and the target in the wind (W2) is

$$\begin{aligned} \mathrm{d}x(t) &= v\cos(\theta)\,\mathrm{d}t + v_w\cos(\theta_w)\,\mathrm{d}t \\ \mathrm{d}y(t) &= v\sin(\theta)\,\mathrm{d}t + v_w\sin(\theta_w)\,\mathrm{d}t \\ \mathrm{d}\theta(t) &= \frac{u}{\rho_{\min}}\,\mathrm{d}t, \qquad |u| \le 1 \\ \mathrm{d}\theta_w(t) &= \sigma_\theta\,\mathrm{d}W_\theta, \end{aligned} \tag{W2}$$

where $\mathrm{d}W_\theta$ is an increment of a Wiener process, and where σ_θ is its corresponding intensity. When $\sigma_\theta = 0$, one obtains a model of constant wind in the direction θ_w. Section 5.4 describes optimal feedback controls for the deterministic case ($\sigma_\theta = 0$) and the stochastic case ($\sigma_\theta > 0$).

The proposed control schemes for the deterministic wind, which are based on ana-

lytic arguments, will give significant insights for the subsequent analysis and will illustrate some interesting patterns of the solution of the stochastic optimal control problem. It will be shown later on that the control strategies for each deterministic wind model, when applied to the DV in the presence of the stochastic wind, will capture the salient features of the solution of the stochastic optimal control problem.

5.3 Feedback Laws with No Wind Information

In this section, feedback control laws are developed that drive the DV to its target in the presence of an unknown wind (W1). First, a method for designing a feedback control for the deterministic problem, that completely ignores the presence of a wind, is briefly discussed. Next, an optimal feedback control will be computed for the case where the Cartesian components of the wind vary stochastically. In all cases, the target set is a ball of radius with $\delta = 0.1$, and the velocity of the vehicle is constant $v = 1$.

5.3.1 Deterministic Case

First, a control law that is completely independent of any information about the distribution of the wind is proposed. In other words, a feedback control law is designed under the assumption that the wind is modeled by (W1) with $\sigma_W = 0$. Therefore, this control law is "blind" to the presence and the statistics of the actual wind. This approach will give two navigation laws that are similar to the pure pursuit strategy from missile guidance [30], which is a control strategy that forces the velocity vector of the

controlled object (the DV in this case) to point towards its destination at every instant of time.

Note that in the presence of a wind and with the application of a feedback law that imitates the pure pursuit strategy, the DV will not be able to instantaneously change its motion in order to point its velocity vector toward the target. This happens for two reasons. The first reason is because the rate at which the DV can rotate its velocity vector is bounded by the turning rate constraint (5.3). The second reason has to do with the fact that, by hypothesis, the pure pursuit law does not account for the wind, and, consequently, even if the DV were able to rotate its forward velocity vector arbitrarily fast, it would be this forward velocity vector that points toward the target rather than the inertial velocity.

Let φ be the angle between the vehicle's forward velocity vector and the line-of-sight to the target, given by $\varphi = \theta - \text{atan}_2(y,x) + \pi$ and mapped to lie in $\varphi \in (-\pi, \pi]$ (see Fig. 5.1). The proposed (suboptimal) pure pursuit-like navigation law takes the following state-feedback form

$$u(\varphi) = \begin{cases} -1 & \text{if } \varphi \in (0, \pi], \\ 0 & \text{if } \varphi = 0, \\ +1 & \text{if } \varphi \in (-\pi, 0). \end{cases} \tag{5.5}$$

One important observation is that the control law (5.5) does not depend on the distance $r(t) = \sqrt{(x(t))^2 + (y(t))^2}$ of the DV from the target but only on the angle φ. The state feedback control law given in (5.5) will be referred to as the *geometric pure pursuit*

(GPP for short) law. Note that the GPP law drives the DV to the line $\mathcal{S}_0 := \{(r,\varphi) : \varphi = 0\}$, which is a "switching surface." In the absence of wind, once the DV reaches $\mathcal{S}_0$, it travels along $\mathcal{S}_0$ until it reaches the target (such that $r = 0$ at the final time T) with the application of the control input $u = 0$. Therefore, the GPP law is a bang-off control law with one switching at most, that is, a control law which is necessarily a control sequence $\{\pm 1, 0\}$.

It is important to highlight that the GPP law turns out to be the time-optimal control law of the MD problem for the majority (but not all) of the initial configurations $[x(0), y(0), \theta(0)]^\intercal$ (see Fig. 5.2), when there is no wind [49, 258]. However, there are still initial configurations from which the DV driven by the navigation law (5.5) either cannot reach the target set at all or can reach the target only suboptimally. The previous two cases are observed, for example, when the DV is close to the target with a relatively large $|\varphi|$.

In particular, it can be shown [49, 258] that if the DV starts, at time $t = 0$, from any point that belongs to one of the two regions, $\mathcal{C}_+$ and $\mathcal{C}_-$, defined by (see Fig. 5.2)

$$\mathcal{C}_- = \{(r,\varphi) \ : \ r \leq 2\rho_{\min}\sin(-\varphi),\ \varphi < 0\} \tag{5.6}$$

$$\mathcal{C}_+ = \{(r,\varphi) \ : \ r \leq 2\rho_{\min}\sin(\varphi),\ \varphi > 0\}, \tag{5.7}$$

then the target cannot be reached by means of the GPP law in the absence of a stochastic wind (scenarios where the stochastic wind helps the DV to reach its target even by means of a GPP law will be shown later on). Therefore, in order to complete the design of a feedback control law for any possible state of the DV, one needs to con-

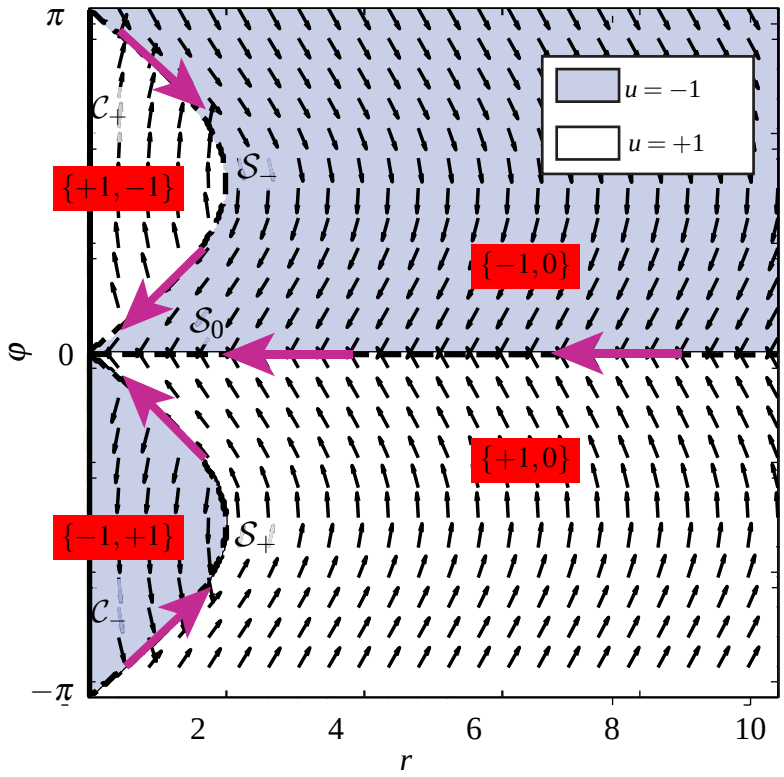


Figure 5.2: Time-optimal partition of the control input space and state feedback control law of the MD problem with free terminal heading in the absence of wind. One can use this control strategy as a feedback law for the case of a stochastic wind. Control sequences for an initial state in each time-optimal partition are indicated in red background.

sider the optimal solution of the MD problem in the case when the terminal heading is free [49, 258]. It turns out that the boundaries of $\mathcal{C}_+$ and $\mathcal{C}_-$, denoted, respectively, by $\mathcal{S}_-$ and $\mathcal{S}_+$ (the choice of the subscript notation will become apparent shortly later), correspond to two new "switching surfaces" along which the DV travels all the way to the target. In particular, when the DV starts in the interior of $\mathcal{C}_+$ (respectively, $\mathcal{C}_-$), then the minimum-time control action is $u = +1$ (respectively, $u = -1$), which may appear to be counterintuitive, since its effect is to increase $|\varphi|$ rather to decrease it. The control input remains constant until the DV reaches the "switching surface" $\mathcal{S}_-$ (respectively, $\mathcal{S}_+$), where the control switches to $u = -1$ (respectively, $u = +1$), and subsequently, the DV travels along $\mathcal{S}_-$ (respectively, $\mathcal{S}_+$) all the way to the target driven by $u = -1$ (respectively, $u = +1$). The net effect is that when the DV starts inside the regions $\mathcal{C}_\pm$, the DV must first distance itself from the target so that its minimum turning radius $\rho_{\min}$ is sufficient to turn towards the target. Note that in this case the control law is bang-bang with one switching at most, that is, a control sequence $\{\pm 1, \mp 1\}$. The situation is illustrated in Fig. 5.2 for $\rho_{\min} = 1$.

The GPP law given in (5.5), therefore, needs to be updated appropriately to account for the previous remarks. In particular, the new feedback control law is given by

$$u(r,\varphi) = \begin{cases} -1 & \text{if } (r,\phi) \in \Sigma_-, \\ 0 & \text{if } (r,\phi) \in \Sigma_0, \\ +1 & \text{if } (r,\phi) \in \Sigma_+, \end{cases} \tag{5.8}$$

where,

$$\Sigma_- := \{(r,\varphi) : \varphi \in (0,\pi]\} \cap (\text{int}\mathcal{C}_+)^c \cup \text{int}\mathcal{C}_-$$

$$\Sigma_+ := \{(r,\varphi) : \varphi \in (-\pi,0)\} \cap (\text{int}\mathcal{C}_-)^c \cup \text{int}\mathcal{C}_+$$

$$\Sigma_0 := \{(r,\varphi) : \varphi = 0\}.$$

The state feedback law (5.8) will be henceforth referred to as the *optimal pure pursuit* (OPP for short) law, because, at every instant of time, it steers the DV to the target based on the optimal strategy that corresponds to its current position. Note that in the absence of wind, the OPP law is the optimal control law of the MD problem with free final heading [49, 258]. One important remark about the OPP law is that the control variable u may attain the value zero, which is in the interior of its admissible set $[-1,1]$. Thus, the control $u = 0$ is *singular*, and the part of the optimal trajectory it generates is referred to as the *singular arc* of the optimal solution. As explained in detail in Ref. [28], singular arcs may be part of the optimal solution of the MD problem in the absence of wind, only when the so-called *switching function* of the constrained optimal control problem, along with its first time derivative, vanish simultaneously (for a non-trivial time interval). Note that while $u = \pm 1$ corresponds to turning left/right, the control $u = 0$ corresponds to straight paths.

Figure 5.3 illustrates the level sets of the minimum time-to-go function, which can be computed analytically by using standard optimal control techniques and geometric tools, as shown in [28].

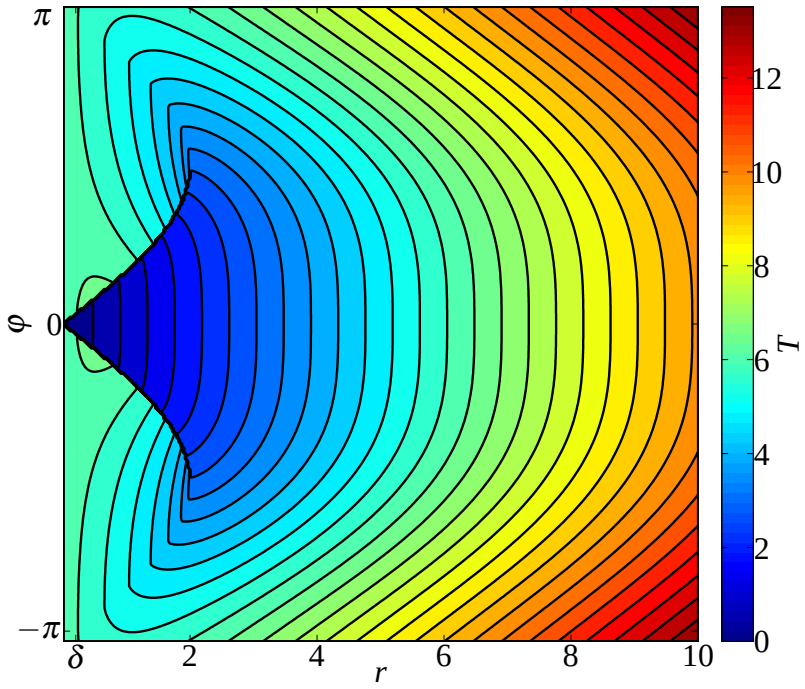


Figure 5.3: Level sets of the minimum time-to-go function of the MD problem with free terminal heading.

5.3.2 Stochastic Case

In this section, an optimal feedback control law for the stochastic kinematic model (W1) and cost functional (5.4) is developed. The optimal control is computed using the *Markov chain approximation method* [137], which ensures that when discretizing a state space for value iteration in stochastic optimal control problems, the chosen spatial and temporal step sizes accurately scale in the same way as in the original stochastic process. The method constructs a discrete-time and discrete-state approximation to the cost function in the form of a controlled Markov chain that is "locally-consistent" with the process under control.

Since the method involves discretization of the state space, one first reduces the

number of dimensions in the model (W1). Applying Itô's differentiation rule to the DV-target distance $r(t)$ and the viewing angle $\varphi(t)$ where $\varphi \in (-\pi, \pi]$, as before (see Fig. 5.1), it can be shown that the relative DV-target system coordinates obey (see the Appendix)

$$\mathrm{d}r(t) = \left(-v\cos(\varphi) + \frac{\sigma_W^2}{2r}\right)\mathrm{d}t + \sigma_W \mathrm{d}W_{\parallel}, \tag{5.9}$$

$$\mathrm{d}\varphi(t) = \left(\frac{v}{r}\sin(\varphi) + \frac{u}{\rho_{\min}}\right)\mathrm{d}t + \frac{\sigma_W}{r}\mathrm{d}W_{\perp}, \tag{5.10}$$

where $|u| \leq 1$, and where $\mathrm{d}W_{\parallel}$ and $\mathrm{d}W_{\perp}$ are mutually independent increments of unit intensity Wiener processes aligned with the direction of DV motion. Note the presence of a positive bias $\sigma_W^2/2r$ in the relation for $r(t)$, which is a consequence of the random process included in the analysis. In the proposed parametrization, only distances $r \geq \delta$ outside the target set are considered, and so (5.9)-(5.10) is well defined.

In the Appendix the equations for value iterations on the cost-to-go function using the Markov Chain approximation method are derived. From this, the optimal angular velocity of the DV may be obtained for any relative distance $r \geq \delta$ and viewing angle φ. The structure of the optimal control law (W1) is seen in Fig. 5.4(a) for $\sigma_W = 0.1$ and discretization steps $\Delta r = 0.02$ and $\Delta\phi = 0.025$. As in the deterministic model ($\sigma_W = 0$) case (Fig. 5.2), the value iteration stationary control law is composed of bang-bang regions instructing the DV to turn left or right and singular arcs. With smaller noise, the optimal control is comprised of four regions, two directing the target to turn left, and others instructing a turn to the right. The reader should note the similarity between Fig. 5.4(a) and the OPP control illustrated in Fig. 5.2. In particular, the structure of the

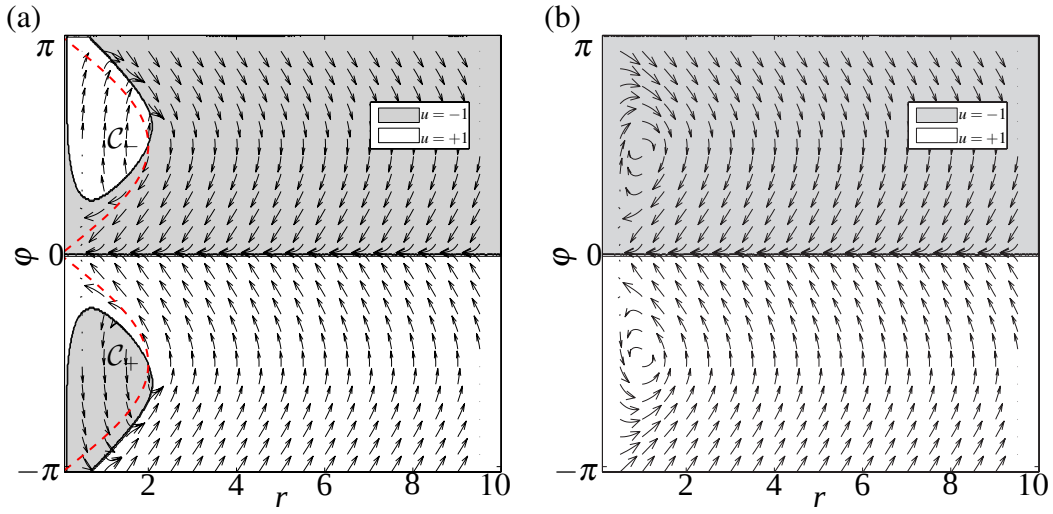


Figure 5.4: Dubins vehicle optimal turning rate control policy $u(r,\varphi)$ for model (W1). (a) Stochastic optimal control policy for $\sigma_W = 0.1$. For comparison, the switching curves from the deterministic model-based OPP control in Fig. 5.2 are outlined in red. (b) Stochastic optimal control policy for $\sigma_W = 0.5$ yields the GPP control (5.5) for deterministic winds.

regions $\mathcal{C}_-$ and $\mathcal{C}_+$ have changed somewhat, as a consequence of the stochastic variation of the wind. In Fig. 5.4(b), a higher noise intensity of $\sigma_W = 0.5$ causes the control to return to GPP control (5.5). In other words, the variance of the process is so large that it becomes exceedingly difficult to predict the relative DV-target state, and the optimal control for the stochastic model matches a simpler, analytically-derived control for the deterministic model that, as described in the previous section, is not optimal for some initial conditions close to the target. This suggests that, for our problem, a deterministic control may suffice for the optimal feedback control when the variance of the stochastic wind is sufficiently large.

This control strategy remains optimal for even larger σ_W, but due to the bias in $r(t)$ (see Eq. (5.9)), this control policy may not be successful in guiding the DV to the

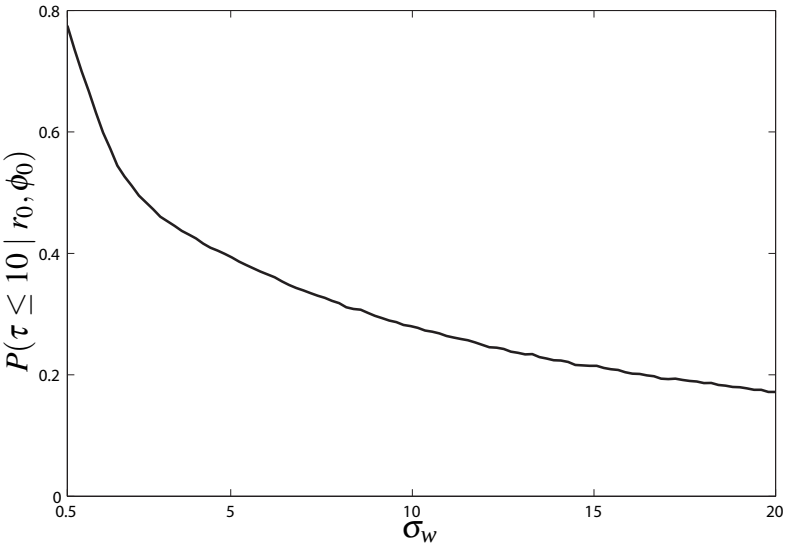


Figure 5.5: Probability of hitting the target set in the time interval $0 < \tau \leq 10$ for an initial condition a distance 1 [m] from the target and facing towards the target ($(r, \varphi) = (1, 0)$).

target in a reasonable amount of time for high values of σ_W. Although a solution to the backward Kolmogorov equation [105] indicates that the DV will eventually hit the target with probability one as $t \to \infty$, the expected value of the hitting time becomes exceedingly large with increasing σ_W. Similarly, one can also consider the probability that the DV, initially located at (r, φ), will hit the target set by a specified time τ as a function of the noise intensity σ_W. Figure 5.5 shows this distribution as computed for $(r, \varphi) = (1, 0)$ and $\tau = 10\,\text{s}$ using 1000 simulations for each σ_W.

5.4 Feedback Laws for Wind at an Angle

Next, the second model (W2), in which the wind is now assumed to take on a direction θ_w with known speed $0 < v_w < 1$, where v_w is constant by hypothesis, is assumed and

the feedback control laws for steering the DV in the presence of this wind are discussed.

5.4.1 Deterministic Case

First, the case when $\sigma_\theta = 0$ and $0 < v_w < 1$ is considered. Note that the fact that $\sigma_\theta = 0$ implies that the direction of the wind becomes constant, and consequently, the wind $w = [w_x,\ w_y]^\mathsf{T}$, where $w_x := v_w \cos\theta_w$ and $w_y := v_w \sin\theta_w$, is a constant vector. Therefore, in this section, it is assumed that the constant wind w is known a priori. The equations of motion of the DV become

$$\begin{aligned} \mathrm{d}x &= v\cos(\theta)\mathrm{d}t + w_x\mathrm{d}t, \\ \mathrm{d}y &= v\sin(\theta)\mathrm{d}t + w_y\mathrm{d}t, \\ \mathrm{d}\theta &= \frac{u}{\rho_{\min}}\mathrm{d}t, \qquad |u| \leq 1. \end{aligned} \tag{5.11}$$

First, a feedback law that is similar in spirit to the GPP law given in Eq. (5.5), which exploits the fact that the wind is known a priori, is designed. In particular, the proposed control law tries to rotate the velocity vector of the DV to point at the target. It is easy to show that the control law (5.5) becomes

$$u(\varphi) = \begin{cases} -1 & \text{if } \psi(\varphi) \in (0, \pi], \\ 0 & \text{if } \psi(\varphi) = 0, \\ +1 & \text{if } \psi(\varphi) \in (-\pi, 0). \end{cases} \tag{5.12}$$

and

$$\psi(\varphi) := \operatorname{atan}_2(\dot{y}, \dot{x}) - \theta + \varphi. \tag{5.13}$$

where ψ, is the angle between the inertial velocity of the DV and the line-of-sight (LOS) as is illustrated in Fig. 5.1 (the angle χ in this figure is equal to $\text{atan}_2(\dot{y},\dot{x})$). As it is shown in [30], the navigation law (5.12) is dual to the so-called *parallel navigation* law from missile guidance. The control law (5.12) is henceforth referred to as the Geometric Parallel Navigation (GPN for short) law.

As mentioned in Section 5.3.1, the GPP law that forces the forward velocity of the vehicle to point towards the target may not always be well defined, especially in the vicinity of the target. The same type of argument applies to the GPN law modulo the replacement of the forward velocity with the inertial velocity. Next, a control law that steers the DV to the target using the optimal control that correspond to the current position of the DV and assuming a constant (e.g., average) wind is presented. This control law is referred to as the Optimal Parallel Navigation (OPN for short) law. Note that similarly to the GPN law, the OPN law does not consider the variations of both the speed and the direction of the wind. By combining the type of arguments used in [49, 258], which deal with the standard MD problem with free terminal heading, along with the analysis presented in [24, 26], one can easily show that the candidate optimal control of the Markov-Dubins problem in the presence of a constant wind corresponds to the four control sequences presented in Section 5.3.1, namely, $\{\pm 1, 0\}$ and $\{\pm 1, \mp 1\}$. The main difference between the solutions of the Markov-Dubins problem in the absence of a wind, which was briefly presented in Section 5.3.1, and the Markov-Dubins problem in the presence of a constant wind, is the switching conditions and, consequently, the

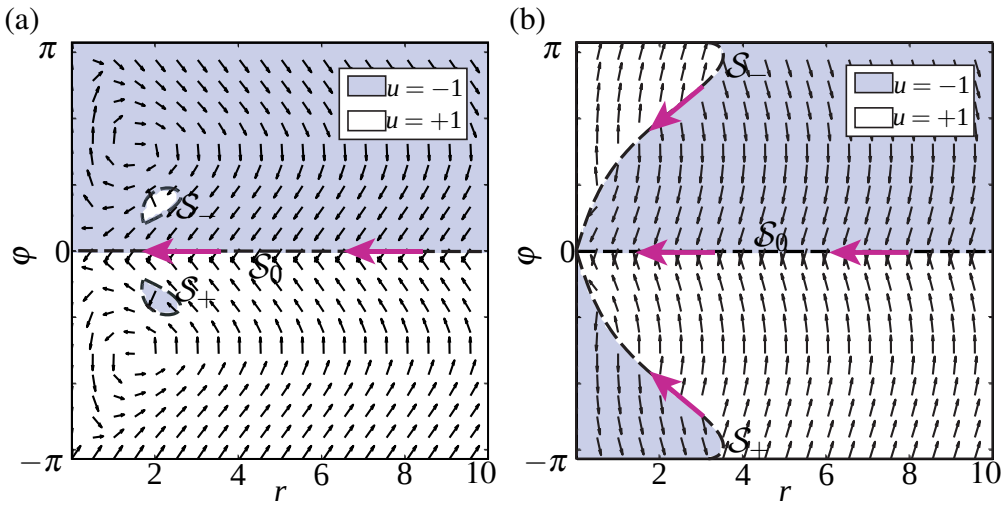


Figure 5.6: Time-optimal partition of the control input space and state feedback control law of the MD problem with free terminal heading in the presence of a constant wind. (a) Tailwind (b) Headwind

switching times of their common control sequences.

Figure 5.6 illustrates the structure of the OPN law in the (r,φ) plane in the presence of a constant tailwind, that is, $\theta_w = \theta(0)$, and a constant headwind, that is, $\theta_w = \pi + \theta(0)$, respectively. One observes that the GPN law coincides with the OPN law for the majority of the boundary conditions especially for the case of a tailwind, whereas in the presence of the headwind the points in the (r,φ) plane where the optimal strategy is bang-bang correspond to a significantly large set. An interesting observation is that the new switching surfaces of the OPN law are associated with those of the OPP law by means of a particular coordinate transformation $\mathcal{H} : (x,y,\theta) \mapsto (x',y',\theta)$, as described in [24]. In particular, a configuration with coordinates (x,y,θ) that belongs to the switching surface $\mathcal{S}_+$, $\mathcal{S}_0$ or $\mathcal{S}_-$ of the OPP law corresponds to a point with coordinates (x',y',θ) that belongs respectively to the switching surface $\mathcal{S}_+$, $\mathcal{S}_0$ and $\mathcal{S}_-$ of the

OPN law, where

$$x' = x + w_x T_{\mathrm{DV}}(x, y, \theta), \tag{5.14}$$

$$y' = y + w_y T_{\mathrm{DV}}(x, y, \theta), \tag{5.15}$$

where $T_{\mathrm{DV}}(x, y, \theta)$ is the minimum time required to drive the DV from (x, y, θ) to the origin with free final heading θ. It is easy to show that for a state $(x, y, \theta) \in \mathcal{S}_+$ $(\mathcal{S}_-)$, it holds that $T_{\mathrm{DV}}(x, y, \theta) = -2\rho_{\min}\varphi(x, y, \theta)/v$ $(2\rho_{\min}\varphi(x, y, \theta)/v)$. In addition, if the state $(x, y, \theta) \in \mathcal{S}_0$, then $T_{\mathrm{DV}}(x, y, \theta) = \sqrt{x^2 + y^2}/v$.

Figure 5.7 illustrates the correspondence of the switching surfaces of the OPN law with those of the OPP law for a tailwind ($\theta_w = \theta(0) = 0$) and a headwind ($\theta_w = \pi + \theta(0) = \pi$) via the previous coordinate transformation. Note that the switching surface $\mathcal{S}_0$ of both the OPG and the OPP laws are the same but the surfaces $\mathcal{S}_\pm$ are different.

Figure 5.8 illustrates the level sets of the minimum time-to-go function in the presence of a constant wind, whose computation entails the solution of a decoupled system of transcendental equations as shown in [24, 26]. In particular, Fig. 5.8(a) and Fig. 5.8(b) illustrate the level sets of the minimum time-to-go function in the presence of a constant tailwind and a constant headwind, respectively.

5.4.2 Stochastic Case

It is now assumed that the direction of the wind θ_w is no longer constant, but is rather described by the stochastic process (W2) with $\sigma_\theta > 0$. A similar derivation to that used

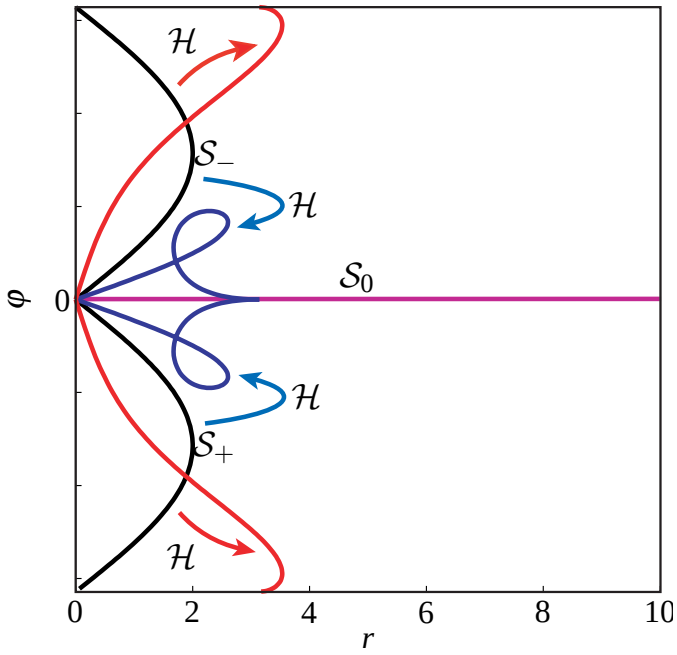


Figure 5.7: Correspondence between the switching surfaces of the OPP and the OPG laws via a transformation for a tailwind (blue curves) and a headwind (red curves).

for model (W1) yields for (W2):

$$\begin{aligned}
\mathrm{d}r(t) &= -\left(v\cos(\varphi) + v_w\cos(\varphi+\gamma)\right)\mathrm{d}t,\\
\mathrm{d}\varphi(t) &= \left(\frac{v}{r}\sin(\varphi) + \frac{v_w}{r}\sin(\varphi+\gamma) + \frac{u}{\rho_{\min}}\right)\mathrm{d}t,\\
\mathrm{d}\gamma(t) &= \frac{u}{\rho_{\min}}\mathrm{d}t - \sigma_\theta \mathrm{d}W_\theta,
\end{aligned} \tag{5.16}$$

where the state $\gamma(t) := \theta(t) - \theta_w(t)$ is introduced to define the difference between the DV heading angle and the direction of the wind θ_w. In the numerical example, the following data are used: $v_w = 0.5$, and $\sigma_\theta = 0.1$. The discretization steps were chosen as $\Delta r = 0.1$, $\Delta\phi = 0.08$, and $\Delta\gamma = 0.12$. As before, value iterations on the optimal cost-to-go were performed as described in the Appendix. Two "slices" of this control, corresponding to the cases where the DV travels in the direction of the wind (tailwind,

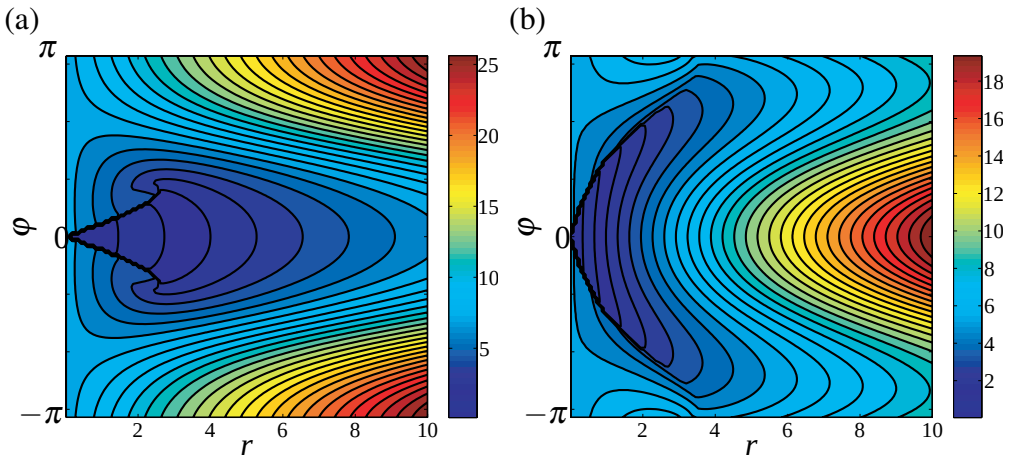


Figure 5.8: Level sets of the minimum time-to-go function of the MD problem with free terminal heading in the presence of a constant wind ($\sigma_\theta = 0$). (a) Tailwind ($\gamma = 0$). (b) Headwind ($\gamma = \pi$).

where $\gamma = 0$) and where it faces the wind (headwind, $\gamma = \pi$) are shown in Fig. 5.9(a) and Fig. 5.9(b), respectively. In each fixed-γ policy, the optimal control resembles that shown in Fig. 5.6, although the location and shape of the switching curves $\mathcal{S}_\pm$ have changed due to the stochastic variation in the wind. In Fig. 5.9(a), only small vestiges of the switching curves are seen, while in Fig. 5.9(b), the shape of these curves has changed. Figure 5.10 shows the expected value of the time required to hit the target in the case of a headwind and tailwind.

5.5 Performance Comparison

In the previous sections, it is seen that the control laws for both deterministic wind models closely resemble their respective optimal feedback control laws for the stochastic wind models. In particular, the control policies for the deterministic and stochastic wind models are identical when far from the target, but differences are seen when r is

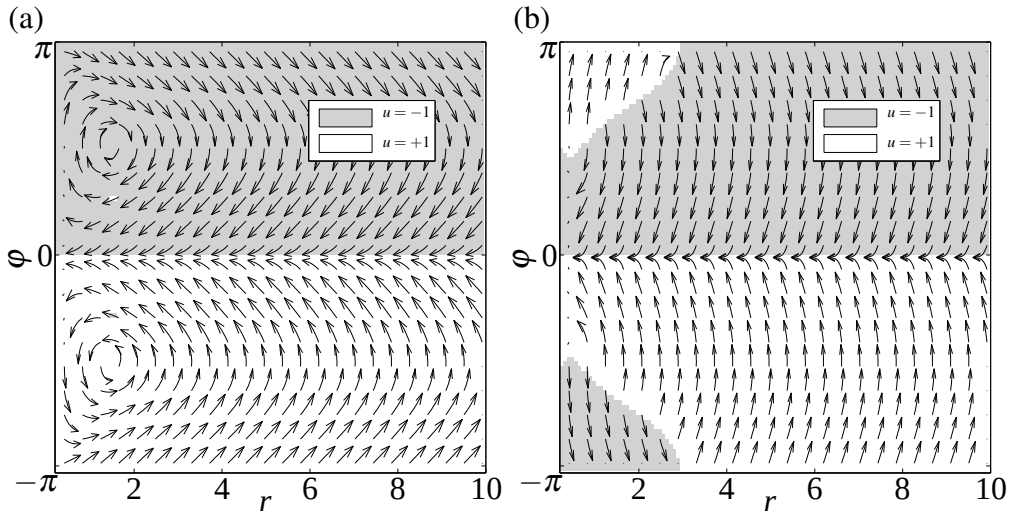


Figure 5.9: Dubins vehicle optimal turning rate control policy $u(r,\varphi,\gamma)$ for stochastic model (W2) with $\sigma_\theta = 0.1$. (a) Tailwind ($\gamma = 0$). (b) Headwind ($\gamma = \pi$).

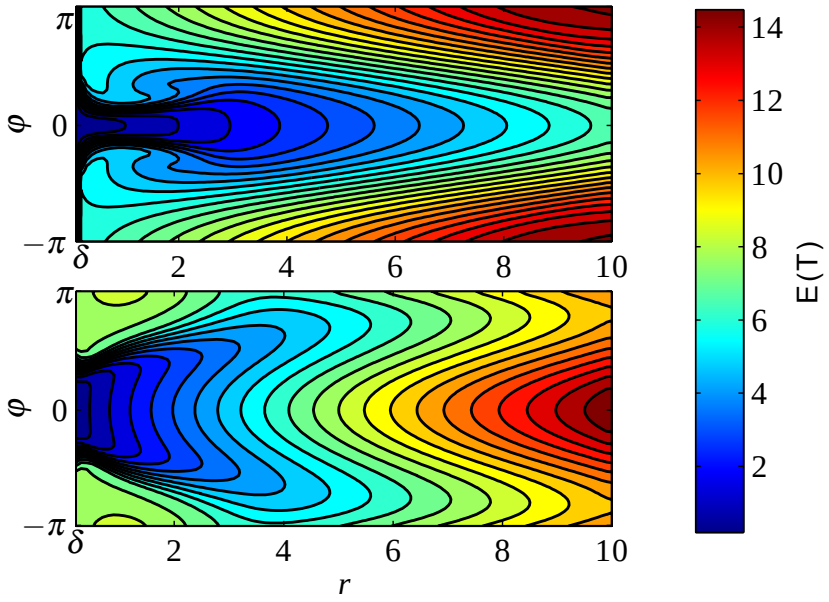


Figure 5.10: Level sets of the minimum time-to-go function of the MD problem with free terminal heading in the presence of a wind with stochastically varying direction and constant speed v_w.

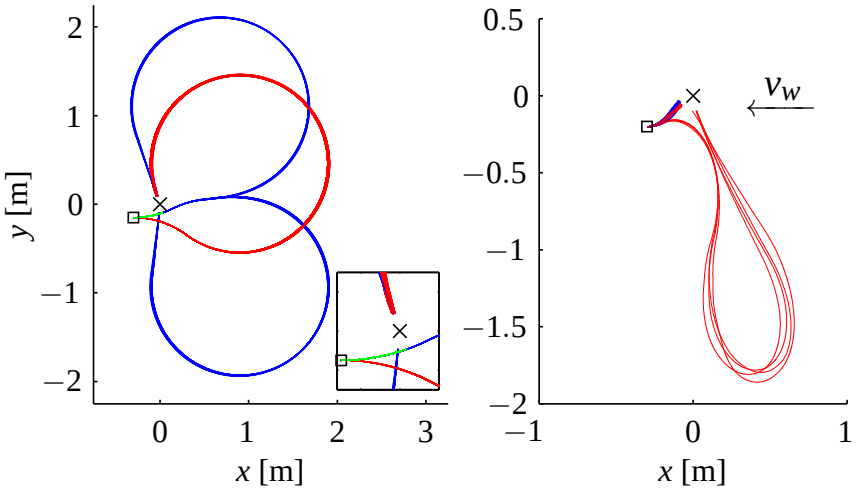


Figure 5.11: DV trajectories for 500 realizations under wind model (W1) (left) and (W2) (right) starting with an initial condition that may result in a difference between trajectories resulting from the control based on the deterministic (red) and stochastic (blue or green) wind models. The initial DV positions are marked with a □, and the target is marked with an ×.

close to δ. To see the effect of these differences, this section provides a comparison of performance of the proposed feedback control laws against the stochastic wind models (W1) and (W2).

As an example, Fig. 5.11 shows a collection of simulated DV trajectories under the controls for the deterministic (red) and stochastic (blue or green) wind models, where the left and right panels correspond to (W1) and (W2), respectively. In this figure, the control anticipating the wind stochasticity assumes that there is a non-zero probability that the stochastic wind may push it beyond its minimum turning radius $\rho_{\min}$ and into the target, and hence the control directs it to perform a left turn. Some realizations (75.6%, shown in green) under this control reach the target, but the remainder (blue) must circle around (see insert). The control for the deterministic wind model directs

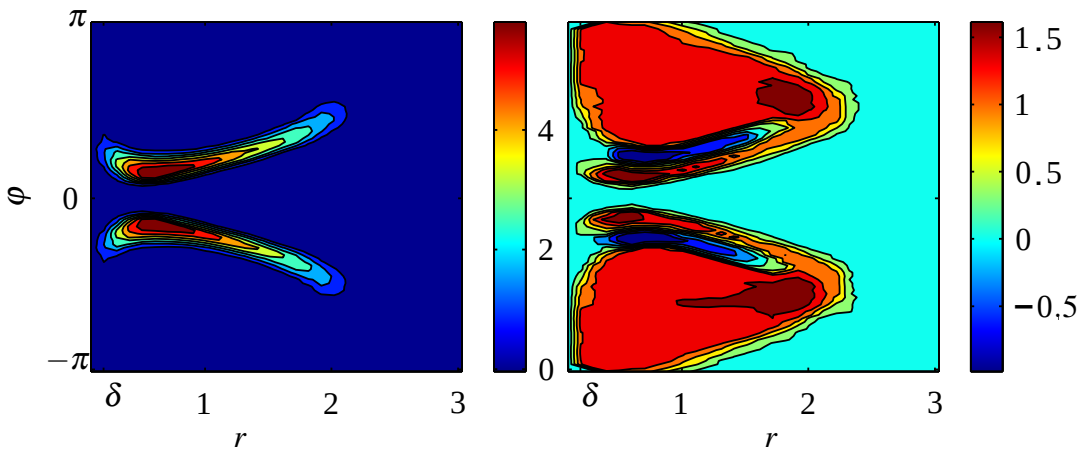


Figure 5.12: Comparison of distributions of time required to hit the target under both OPP control (5.8) and the stochastic optimal control $u(r,\varphi)$, $\sigma_W = 0.1$ in the presence of the stochastic wind (W1). Left: difference in mean hitting time $\mathbf{E}(T_{\text{OPP}}) - \mathbf{E}(T_{\text{stoch}})$. The OPP control results higher $\mathbf{E}(T)$ in regions where the stochastic model-based control differs from the OPP control. Right: difference in standard deviation $\text{std}(T_{\text{OPP}}) - \text{std}(T_{\text{stoch}})$. As expected, the OPP control results in higher standard deviation in the majority of cases.

the red DVs to first distance themselves before approaching the target. Consequently, the regions in the (r,φ) state space corresponding to the trajectories in this example lead to a smaller expected time to hit the target for the stochastic model-based control, as seen in Fig. 5.12. However, there is also a chance that the stochastic model-based control is unsuccessful in hitting the target on its first pass, and so the DV must circle around again. In other words, the stochastic model-based control "risks" a turn toward the target for small r and small φ. Although the expected value of the hitting time decreases under the control anticipating the stochastic winds, the standard deviation of these times may simultaneously increase, as seen in Fig 5.12.

In the right panel of Fig. 5.11, a similar result is seen for the case of wind at an angle (indicated by a v_w arrow). In this case, a small number of the realizations for

the DVs under the deterministic model-based control are affected by the changing wind and must take a longer route to reach the target, whereas the DVs under the stochastic model-based control anticipate the changing wind direction. Similarly, Fig. 5.13 shows the mean time-to-go under (W2) using both the control for the deterministic model shown in Fig. 5.6 and the control law for the stochastic wind model in Fig. 5.9. As before, the expected time-to-go is larger for the deterministic model-based control in regions where the control laws differ. However, unlike (W1), the standard deviation under the stochastic model-based control was consistently smaller since the control accounts for the stochastic wind without instructing for a potentially "risky" approach to the target.

5.6 Conclusions and Future Work

In this paper, the problem of guiding a vehicle with Dubins-type kinematics to a prescribed target set with free final heading in the presence of a stochastic wind in minimum expected time has been addressed. Two approaches to this problem have been proposed. The first one, which was based on analytic techniques, was to employ feedback control laws, based on a deterministic model, that are similar to the well-studied pure pursuit and the parallel navigation laws from the field of missile guidance. The proposed feedback control laws are time-optimal in the absence of wind or in the presence of a wind that is constant and known a priori.

The second approach was to tackle the problem computationally by employing nu-

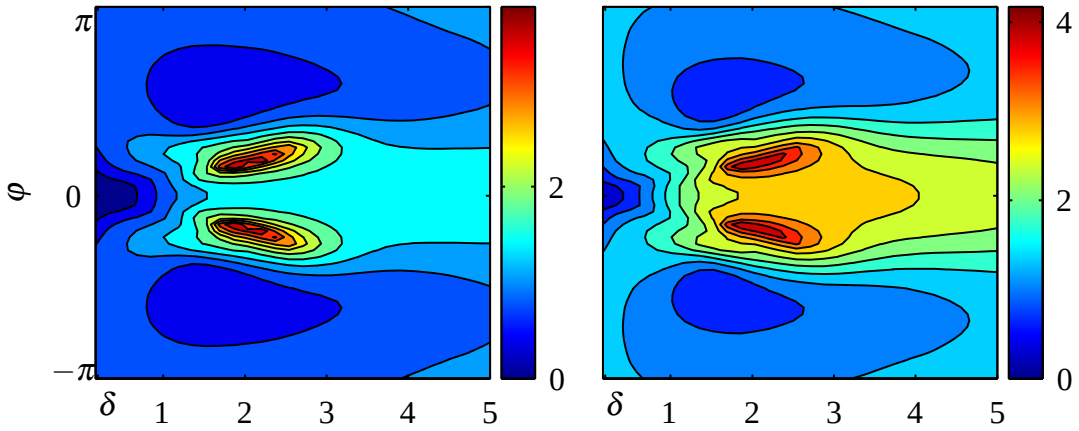
merical tools from stochastic optimal control theory. Because these control laws are based explicitly on the stochastic wind models, they anticipate the wind stochasticity, and the time necessary to steer the Dubins vehicle to the target set in the presence of a stochastic wind is, on average, lower than that under the control for the corresponding deterministic model. However, although the feedback control laws for the deterministic model become suboptimal in the presence of a stochastic wind, it turns out that they still manage to steer the Dubins vehicle to its target set with an acceptable miss target error. On the other hand, a stochastic framework leads to higher expected precision in terms of target miss-distance and more predictable trajectories.

The fact that the deterministic model-based controls perform so well for this problem even in the presence of an unknown stochastic wind is mainly owing to the fact that they are in a feedback form, thus providing a certain degree of robustness against uncertainties. Having that in mind, it may not be surprising that the presented deterministic model-based control laws can work in the presence of small stochastic disturbances, although non-optimally. This may not be the case for other problems in practice where one is only able to generate reliable deterministic *open-loop* trajectories. Surprisingly perhaps, the computation of optimal feedback controls based on *stochastic* models generally is no more difficult (or even easier) than for their deterministic counterparts as the latter can be consistently discretized and cast as a controlled Markov decision process, as shown in this paper. On the other hand, the closed-form feedback laws based on the deterministic model presented in this paper may be more appealing than their

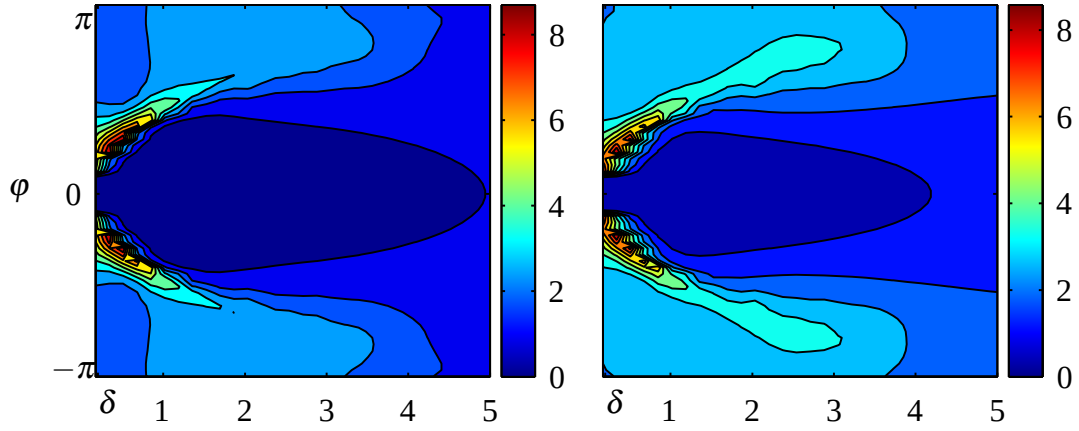
stochastic model-based counterparts, owing to their ease of implementation.

Thus, the similarity between control policies under different levels of wind stochasticity seems to support the use of the feedback controls for deterministic wind models in lieu of stochastic model-based feedback controls when the stochastic effects are small, or can be used as "seeds" that may expedite the computation of the solution to the stochastic optimal control problem, or aid in the verification of numerical results. Moreover, since the role of noise in designing feedback control policies is not fully understood, a side-by-side comparison of the feedback laws for deterministic and stochastic models in other problems may provide useful insights toward a more general theory.

Future work will include the extension of the techniques presented herein to problems with a more realistic model of the wind, including wind models that depend explicitly on the position of the Dubins vehicle. Another possible extension is to characterize control laws for stochastic wind models that minimize a cost function taking into consideration both the expected value and the variance of the time-to-go.



(a) Tailwind ($\gamma = 0$).



(b) Headwind ($\gamma = \pi$).

Figure 5.13: Comparison of distributions of time required to hit the target under both OPP control (Fig. 5.6) and the stochastic optimal control (Fig. 5.9) in the presence of the stochastic wind (W2). Left: difference in mean hitting time $\mathbf{E}(T_{\text{deter}}) - \mathbf{E}(T_{\text{stoch}})$. Right: difference in standard deviation $\text{std}(T_{\text{deter}}) - \text{std}(T_{\text{stoch}})$.

Chapter 6

Stochastic Optimal Enhancement of Distributed Formation Control Using Kalman Smoothers

This chapter is a preprint to the paper

- Anderson, R. P., and Milutinović, D., "Stochastic Optimal Enhancement of Distributed Formation Control Using Kalman Smoothers", Accepted for publication in *Robotica*.

6.1 Introduction

The focus of this work is on an approach to optimally enhance a given distributed feedback control for nonholonomic agents, such as mobile robots [190], or vehicles. We are motivated by the problem of distributed formation control, in which each agent is tasked with attaining and maintaining pre-specified distances from its neighbours. Various control approaches have been proposed for formation control, including PID control [8, 198], artificial potentials [87, 192, 252], navigation functions [211], constraint forces [297], geometric methods [57], adaptive control design [220, 233], sliding mode

control [102], and consensus algorithms[74, 206], for example. These approaches are attractive due to their often intuitive design principles, the transparent, analytic forms of the resulting control laws, and their generally satisfactory results. However, developing a control law that is both optimal and robust to uncertainty presents a greater challenge.

Here, we begin with a *reference* distributed control policy [252] for formations of nonholonomic vehicles consisting of a feedback-controlled turning rate and acceleration and based on an artificial potential function. Next, we numerically compute the additional additive control input necessary to drive the non-optimal system into a formation *optimally* and in a manner that is *robust to uncertainty*. Consequently, our control approach is optimal, and, due to the adopted artificial potential function [252], it provides collision-free agent trajectories. Moreover, in some sense to be defined later, the computed optimal feedback control is “close" to the reference feedback control, preserving many of the nicer properties of the latter, including collision avoidance and, to some extent, qualitative behaviour of the controlled agents. We are not aware of other works that improve upon existing, rather than build from the ground up, mobile robot feedback control laws in this manner.

To compute an optimal feedback control, one must solve the Hamilton-Jacobi-Bellman (HJB) equation, which is a nonlinear partial differential equation (PDE), and, accordingly, the solution of the optimal control problem is necessarily numerical. However, the computational complexity of the solution to the HJB equation grows exponentially with the state space dimension of the system (i.e., with the size of the robotic

team), making conventional approaches to computing the stochastic optimal feedback control intractable. In this work, we exploit the distributed nature of the problem at hand in order to make the solution to the HJB equation computationally feasible. The distributed formation control problem is inherently stochastic – from the perspective of one agent, its neighbours' observations and control inputs are unknown as well as the effects of these inputs on agent trajectories due to model uncertainties. Accordingly, this work considers the problem of controlling one agent based on its own observations of its neighbours in a way that anticipates the probability distribution of their future motion. This probability distribution arises from an assumption that a prior for the unknown control input of an agent can be robustly described as Brownian motion [274], which, for the agent model considered in this paper, results in a so-called "banana distribution" prior [154]. Based on our prior and the system kinematics, we can induce a probability distribution of the relative state $\mathbf{x}$ to all neighbours in an interval $(\mathbf{x}, \mathbf{x} + \mathrm{d}\mathbf{x})$ at a particular future time [280]. It follows that the cost function to be minimized by an agent not only computes the optimal control with respect to the current system state as viewed by that agent, but also with respect to the distribution of possible trajectories originating from the current state.

Besides aiding in the creation of a robust control law, this distribution over future system trajectories serves several additional purposes. For example, the distribution encodes the same type of information that must be repeatedly transmitted among neighbours in some other distributed formation control approaches [84], e.g., assumed future

trajectories of a neighbour. Perhaps more importantly, this distribution over future system trajectories can be used to statistically infer the probability distribution of the *control*, and, hence, the optimal additive control required for the considered reference feedback control. In particular, the relations between the solutions to optimal control PDEs and the probability distribution of stochastic differential equations [101, 183, 292] allows certain stochastic optimal control problems to be written as an estimation problem on the distribution of optimal trajectories in continuous state space in a manner known as the path integral (PI) approach [124, 125, 127, 263, 264].[1]

Related works incorporating the PI framework for multi-agent systems [271, 272, 285, 286] designed control for systems in which agents cooperatively compute their control from a marginalisation of the joint probability distribution of the group's trajectory. In this article[2], we develop a method by which agents independently compute their controls without explicit communication. Moreover, previous works using the PI approach have formulated an unconstrained receding horizon optimal control problem for which stability is difficult to guarantee [122]. We therefore consider two formulations admitting time-invariant feedback control policies. The first is based on a planning horizon that ends only when the formation is reached, while the other is based on an infinite-horizon, discounted cost control problem.

Additionally, we establish a connection between the presented optimal feedback control problem and nonlinear Kalman smoothing algorithms, so that each agent can

[1] There is also an analogous approach in the open-loop control case [169, 185].

[2] Preliminary versions of some portions of this work have previously been presented elsewhere [12, 14].

compute its control in real-time in a way that preserves the optimality and stability properties of the HJB equation solution. That a Kalman smoother can be used to compute an optimal control is related to the well-established duality between linear-quadratic-Gaussian (LQG) control and linear-Gaussian estimation [239]. Although other robotic control algorithms have employed Kalman filters or smoothers for motion planning and path planning [53, 98, 140, 269, 270], the relation of the nonlinear Kalman smoothing algorithm to the HJB equation has not been exploited. Moreover, since our Kalman smoothing control computations are based on the current system state and do not require a global HJB equation solution, we describe how our algorithm may be used to test, pointwise in state space, the possibility for analytic improvements to the reference feedback control for optimality and robustness to uncertainties.

In our approach, after each agent computes and applies the optimal control computed by its Kalman smoother, it observes the new state of its neighbours, and then the process repeats. This method is similar in spirit to receding horizon control/model predictive control (MPC) [161] and related suboptimal algorithms [44], some of which have been used previously for vehicle formation control [62, 84, 174]. However, our approach differs in three respects. First, as previously discussed, the control solution is derived from the HJB equation instead of an open-loop, numerical optimization problem. In addition, for reasons that will become clear in the sequel, the system trajectories to which the Kalman smoother is applied are uncontrolled, whereas model predictive control treats the control along these paths as decision variables. Finally, our Kalman

smoother solution is generally computationally fast compared to nonlinear optimization methods for systems with large dimensional state and control space, and the algorithm quite naturally handles stochasticity. We point out that without the notion of uncertainty in a neighbour's control, an optimization problem resulting from distributed model predictive control may predict a neighbour's trajectory incorrectly and without an explicit remedy for this issue.

This article is organized as follows. Section 6.2 introduces the formation control problem as viewed by a single agent in the group, followed by a derivation of a path integral representation in Section 6.3. Section 6.4 presents a Kalman smoother method for computing individual agent control. Section 6.5 compares the optimal feedback control computed by Kalman smoothers against that computed using a numerical method involving discretisation of the state space and examines the areas of the state space where the analytical reference feedback control could be improved. We continue in Section 6.5 to illustrate our method with simulations of five agents achieving formations, and we conclude with Section 6.6.

6.2 Control Problem Formulation

We consider a team of agents, each described by a Cartesian position (x_m, y_m), a heading angle θ_m, a speed v_m, $m = 1, \ldots, M_{\text{tot}}$, and the kinematic model:

$$\begin{aligned} \mathrm{d}x_m(t) &= v_m \cos\theta_m \mathrm{d}t \\ \mathrm{d}y_m(t) &= v_m \sin\theta_m \mathrm{d}t \\ \mathrm{d}\theta_m(t) &= \omega_m \mathrm{d}t + \sigma_{\theta,m} \mathrm{d}w_{\theta,m} \\ \mathrm{d}v_m(t) &= a_m \mathrm{d}t + \sigma_{v,m} \mathrm{d}w_{v,m}, \end{aligned} \tag{6.1}$$

where ω_m and a_m are the feedback-controlled turning rate and acceleration, respectively, and where $\mathrm{d}w_{\theta,m}$ and $\mathrm{d}w_{v,m}$ are mutually independent Wiener process increments with corresponding intensities $\sigma_{\theta,m}$ and $\sigma_{v,m}$, respectively. Our goal is for the distances separating the agents to reach a set of predefined nominal distances δ_{mn}, $m, n = 1 \ldots, M_{\text{tot}}$, $m \neq n$, and for the heading angles and speeds to be equal.

To achieve a distributed control, the problem is formulated from the perspective of just one agent whose state (x, y, θ, v) is notated without subscript. This agent observes M neighbours with subscripts $m = 1, \ldots, M$, $M \leq M_{\text{tot}}$, regardless of the total number of agents M_{tot} in the team, and computes an optimal feedback control to reach a formation with respect to its observed neighbours. No further information is assumed about the observations made by its neighbours. In other words, the only interactions modeled by an agent are undirected (bilateral) and only include that agent's observed neighbours. This is illustrated by Figure 6.1.

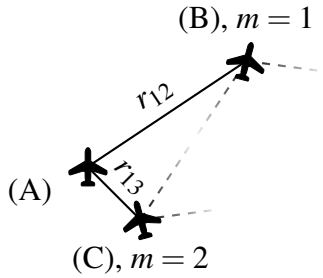


Figure 6.1: In this scenario, agent A observes neighbours B and C, labeled by agent A as $m = 1$ and $m = 2$, and attempts to achieve the inter-agent spacings $r_{12} = \delta_{12}$ and $r_{13} = \delta_{13}$, as well as alignment of heading angles and speeds. Agent A is unaware of any other observation (dashed lines). Consequently, both the reference control and the optimal controls computed by agent A do not include the observation B–C.

We introduce collision avoidance by adding to the original kinematic model an artificial potential function-based control [252] for collision-free velocity vector alignment of groups of vehicles described by a noiseless version of (6.1). The evolution equations for $\theta(t)$ and $v(t)$ become

$$\mathrm{d}\theta(t) = \omega_R \mathrm{d}t + \omega \mathrm{d}t + \sigma_\theta \mathrm{d}w_\theta \tag{6.2}$$

$$\mathrm{d}v(t) = a_R \mathrm{d}t + u \mathrm{d}t + \sigma_v \mathrm{d}w_v, \tag{6.3}$$

so that the controls ω and a are interpreted as the optimal *correction* terms to the reference turning rate feedback control ω_R and acceleration feedback control a_R, respectively, in the presence of uncertainty. For brevity, we omit details [252] for the equations for ω_R and a_R. Suffice it to say that in the deterministic case ($\sigma_{\theta,m} = \sigma_{v,m} = 0$), the reference feedback control ω_R and a_R ensures collision avoidance, that is, the inter-agent distances remain strictly positive[3], and it aligns the agent velocity vectors. It

[3]Ensuring collision avoidance among agents with non-zero collision radii would require a different deterministic control or artificial potential function than is considered in this work.

guarantees that the group will tend toward a minimum of the an artificially-constructed potential $V(r_m)$, where $r_m = \sqrt{(x-x_m)^2 + (y-y_m)^2}$, $m = 1, \ldots, M$. Our specific choice of $V(r_m) = \delta_m^2 ||r_m||^{-2} + 2\log||r_m||$ causes the potential [252] to reach minimum value when all inter-agent distances $r_m \to \delta_m$.

Evolutions of the heading angle $\theta_m(t)$ and speed $v_m(t)$ of a neighbour m are based on its own observations that are unknown to any other agent. Therefore, their increments for an agent m are modeled as Gaussian random variables:

$$\mathrm{d}\theta_m(t) = \omega_{R,m}\mathrm{d}t + N\left(\omega_m \mathrm{d}t, \sigma^2_{\theta,m}\mathrm{d}t\right) \tag{6.4}$$

$$\mathrm{d}v_m(t) = a_{R,m}\mathrm{d}t + N\left(a_m \mathrm{d}t, \sigma^2_{v,m}\mathrm{d}t\right), \tag{6.5}$$

where $\omega_{R,m}$ and $a_{R,m}$ are the reference controls computed for neighbouring agents (see Figure 6.1 caption), and where $\sigma_{\theta,m}$ and $\sigma_{v,m}$ take into account both kinematic uncertainty *and control uncertainty*, so that $\sigma_{\theta,m} \geq \sigma_\theta$ and $\sigma_{v,m} \geq \sigma_v$. In summary, the kinematic model for one agent with M observed neighbours takes the form

$$\mathrm{d}\Delta x_m(t) = \mathrm{d}(x(t) - x_m(t)) = v\cos\theta \mathrm{d}t - v\cos\theta_m \mathrm{d}t \tag{6.6}$$

$$\mathrm{d}\Delta y_m(t) = \mathrm{d}(y(t) - y_m(t)) = v\sin\theta \mathrm{d}t - v\sin\theta_m \mathrm{d}t \tag{6.7}$$

$$\mathrm{d}\theta(t) = \omega_R \mathrm{d}t + \omega \mathrm{d}t + \sigma_\theta \mathrm{d}w_\theta \tag{6.8}$$

$$\mathrm{d}v(t) = a_R \mathrm{d}t + a\mathrm{d}t + \sigma_v \mathrm{d}w_v \tag{6.9}$$

$$\mathrm{d}\theta_m(t) = \omega_{R,m}\mathrm{d}t + \omega_m \mathrm{d}t + \sigma_{\theta,m}\mathrm{d}w_{\theta,m} \tag{6.10}$$

$$\mathrm{d}v_m(t) = a_{R,m}\mathrm{d}t + a_m \mathrm{d}t + \sigma_{v,m}\mathrm{d}w_{v,m}, \qquad m = 1, \ldots, M,\ M \leq M_{\mathrm{tot}}. \tag{6.11}$$

This model can be written in a general form:

$$\mathrm{d}\mathbf{x}(t) = f(\mathbf{x})\mathrm{d}t + B\mathbf{u}\mathrm{d}t + \Gamma\mathrm{d}\mathbf{w}, \tag{6.12}$$

where the state vector $\mathbf{x}$ includes the system state from the perspective of just one agent, $f(\mathbf{x})$ captures the kinematics including the deterministic, collision-avoiding reference controls, and $\mathbf{u} = [\omega, a, \omega_1, a_1, \ldots, \omega_M, a_M]^\intercal$ is a vector of optimal feedback controls to be computed by the agent. The Wiener process $\mathrm{d}\mathbf{w}$ captures the uncertainty due to model kinematics for each agent, as well as the uncertainty due to the control executed by neighbouring agents $m = 1, \ldots, M$.

We consider two types of cost functionals that aim to bring the agents into formation. In the first, our goal is to compute the feedback controls $\mathbf{u}(\mathbf{x})$ that minimize the total accumulated cost up until the formation is reached (a problem sometimes called control until a target set is reached). The formation is achieved when the inter-agent distances $r_m = \sqrt{(\Delta x_m)^2 + (\Delta y_m)^2}$, $m = 1, \ldots, M$, reach a set of predefined nominal distances δ_m, the agents' heading angles are aligned, and the speeds are equal. Since this is a stochastic problem, we must define the target set $\mathbf{x} \in \mathcal{F}$ as a small ball about the formation. We define the following cost functional:

$$J(\mathbf{x}) = \min_{\mathbf{u}} \mathbb{E}\left\{\int_0^\tau \frac{1}{2}\Big(k(\mathbf{x}) + \mathbf{u}^\intercal R\mathbf{u}\Big)\mathrm{d}s\right\}, \tag{6.13}$$

where $\tau = \inf\{t > 0 : \mathbf{x}(t) \in \mathcal{F}\}$ is a (finite) first exit time, i.e., the first time that the state reaches the formation $\mathcal{F}$. We note that, unlike previous works that either use a receding horizon approach or fix a final time, here the final time τ is not known in

advance. The positive semi-definite matrix R in (6.13) provides a quadratic control penalty, and the instantaneous state cost $k(\mathbf{x})$,

$$k(\mathbf{x}) = (h(\mathbf{x}) - \boldsymbol{\mu})^{\mathsf{T}} Q (h(\mathbf{x}) - \boldsymbol{\mu}), \tag{6.14}$$

is a quadratic that reaches minimum value when the inter-agent distances r_m equal the predefined nominal distances δ_m, the heading angles are equal, and the speeds are equal:

$$h(\mathbf{x}) = [r_1, \ldots, r_M, \theta - \theta_1, \ldots, \theta - \theta_M, v - v_1, \ldots, v - v_M]^{\mathsf{T}} \tag{6.15}$$

$$\boldsymbol{\mu} = [\delta_1, \ldots, \delta_M, 0, \ldots, 0]^{\mathsf{T}} \tag{6.16}$$

and where Q is a diagonal positive definite matrix. The form of the cost function relieves the agents' dependence on a global coordinate frame. Since all couplings between pairs of agents in the cost function are relative, the coordinate frame implied by the Cartesian components Δx_m and Δy_m and angles θ and θ_m in (6.6)-(6.11) do not need to be shared among the agents. Further note that this instantaneous state cost (6.14) reaches minimum value when the potential $V(r_m)$ associated with the reference control also reaches minimum value. However, the reference control also prevents collisions using an infinite potential energy when inter-agent distances approach $r_m = 0$. Since the correction control $\mathbf{u}$ is penalized, it will not overcome this barrier, and collision avoidance is still ensured.

The second type of control problem we consider is to minimize an infinite horizon

cost functional with a discounting factor $\beta > 0$:

$$J(\mathbf{x}) = \min_{\mathbf{u}} \mathbb{E}\left\{\int_0^\infty \frac{e^{-\beta s}}{2}\Big(k(\mathbf{x}) + \mathbf{u}^\mathsf{T} R\mathbf{u}\Big)\mathrm{d}s\right\}, \tag{6.17}$$

where the terms Q, R, and $k(\mathbf{x})$ are as in (6.13).

6.3 Path Integral Representation

In this section we show how the optimal control problem can be represented as a path integral over possible system trajectories. The derivation is similar to that used in previous works [271], but the new types of cost functionals used in this paper warrant a new derivation. We will first consider the cost functional for control until the formation is reached (6.13), hereinafter abbreviated CUF.

6.3.1 Control until formation (CUF)

The (stochastic) Hamilton-Jacobi-Bellman equation for the model (6.12) and cost functional (6.13) is

$$0 = \min_{\mathbf{u}}\left\{(f + B\mathbf{u})^\mathsf{T}\partial_{\mathbf{x}} J + \frac{1}{2}\mathrm{Tr}\left(\Sigma\partial_{\mathbf{x}}^2 J\right) + \frac{1}{2}k(\mathbf{x}) + \frac{1}{2}\mathbf{u}^\mathsf{T} R\mathbf{u}\right\}, \tag{6.18}$$

where $\Sigma = \Gamma\Gamma^\mathsf{T}$ and the boundary condition for this PDE as

$$J(\mathbf{x}(\tau)) = 0, \qquad \mathbf{x} \in \mathcal{F}. \tag{6.19}$$

The HJB equation must typically be solved numerically in a discretised state space until a steady state is reached (see [137], for example). However, the structure of the problem

at hand allows us to avoid the discretisation through a suitable transformation.

The optimal control $\mathbf{u}(\mathbf{x})$ that minimizes (6.18) is

$$\mathbf{u}(\mathbf{x}) = -R^{-1}B^{\mathsf{T}}\partial_{\mathbf{x}}J, \tag{6.20}$$

which, when substituted back into the HJB equation, yields:

$$0 = f^{\mathsf{T}}\partial_{\mathbf{x}}J - \frac{1}{2}(\partial_{\mathbf{x}}J)^{\mathsf{T}}BR^{-1}B^{\mathsf{T}}\partial_{\mathbf{x}}J + \frac{1}{2}\mathrm{Tr}\left(\Sigma\partial_{\mathbf{x}}^2 J\right) + \frac{1}{2}k(\mathbf{x}(t)). \tag{6.21}$$

Next, we apply a logarithmic transformation [96] $J(\mathbf{x}) = -\lambda\log\Psi(\mathbf{x})$ for constant $\lambda > 0$ to obtain a new PDE

$$\begin{aligned} 0 = \frac{k(\mathbf{x})}{2\lambda} - \frac{f^{\mathsf{T}}}{\Psi}\partial_{\mathbf{x}}\Psi - \frac{1}{2}\frac{1}{\Psi}\mathrm{Tr}\left(\Sigma\partial_{\mathbf{x}}^2\Psi\right) \\ - \frac{1}{2}\frac{\lambda}{\Psi^2}(\partial_{\mathbf{x}}\Psi)^{\mathsf{T}}BR^{-1}B^{\mathsf{T}}\partial_{\mathbf{x}}\Psi + \frac{1}{2}\frac{1}{\Psi^2}(\partial_{\mathbf{x}}\Psi)^{\mathsf{T}}\Sigma\partial_{\mathbf{x}}\Psi. \end{aligned} \tag{6.22}$$

In the model (6.6)-(6.11), it can be seen that the optimal controls $\mathbf{u}(\mathbf{x})$ act as a correction term to the deterministic controls and the stochastic noise. Penalizing this control (6.13) suggests that the optimal control is that which is "close" (in terms of Kullback-Leibler divergence [264]) to the reference control. Moreover, this implies that the possibility of a large stochastic disturbance (either due to neighbours' unknown controls or model kinematics) requires the possibility of a greater control input. Because of this, we assume that the noise in the controlled components is inversely proportional to the control penalty, or

$$\Sigma = \Gamma\Gamma^{\mathsf{T}} = \lambda BR^{-1}B^{\mathsf{T}}. \tag{6.23}$$

This selects the value of the control penalty that we shall use in the sequel as

$$R = \lambda \, diag\left(\sigma_\theta^{-2}, \sigma_v^{-2}, \sigma_{\theta,1}^{-2}, \sigma_{v,1}^{-2}, \ldots, \sigma_{\theta,M}^{-2}, \sigma_{v,M}^{-2}\right) \tag{6.24}$$

and also causes the quadratic terms on the second line of (6.22) to cancel, so that the remaining PDE for Ψ is linear:

$$0 = f^\intercal \partial_{\mathbf{x}} \Psi(\mathbf{x}) + \frac{1}{2}\mathrm{Tr}\left(\Sigma \partial_{\mathbf{x}}^2 \Psi\right) - \frac{k(\mathbf{x})}{2\lambda}\Psi(\mathbf{x}), \tag{6.25}$$

$$\Psi(\mathbf{x}) = 1, \qquad \mathbf{x} \in \mathcal{F}. \tag{6.26}$$

As before, this could be solved numerically until a steady state is reached. However, the Feynman-Kac equations [183, 292] connect certain linear differential operators to adjoint operators that describe the evolution of a *forward* diffusion process beginning from the current state $\widetilde{\mathbf{x}}(0) = \widetilde{\mathbf{x}}_0 = \mathbf{x}$. From the Feynman-Kac equations, the solution to (6.25) is [101]:

$$\Psi(\mathbf{x}) = \mathbb{E}_{\widetilde{\mathbf{x}},\tau|\widetilde{\mathbf{x}}_0}\left\{\Psi(\widetilde{\mathbf{x}}(\tau))\exp\left(-\frac{1}{2\lambda}\int_0^\tau k(\widetilde{\mathbf{x}}(s))ds\right)\right\}. \tag{6.27}$$

where $\widetilde{\mathbf{x}}(t)$ satisfies the path integral-associated, uncontrolled dynamics (cf. (6.12)),

$$\mathrm{d}\widetilde{\mathbf{x}}(t) = f(\widetilde{\mathbf{x}}(t))\mathrm{d}t + \Gamma \mathrm{d}\mathbf{w}, \tag{6.28}$$

with initial condition $\widetilde{\mathbf{x}}(0) = \mathbf{x}$, and which, as before, includes the reference control inputs for one agent and its observed neighbours (see Fig. 6.1). The expectation in (6.27) is taken with respect to the joint distribution of $(\widetilde{\mathbf{x}}, \tau)$ of sample paths $\widetilde{\mathbf{x}} = \widetilde{\mathbf{x}}(t)$ that begin at $\widetilde{\mathbf{x}}_0 = \mathbf{x}$ and evolve as (6.28) until hitting the formation $\widetilde{\mathbf{x}}(\tau) \in \mathcal{F}$ at time τ. Unlike previous PI works, where the terminal time is fixed and known in advance, this

stopping time is a property of the set of stochastic trajectories $\widetilde{\mathbf{x}}(t)$.

The distribution $(\widetilde{\mathbf{x}}, \tau)$ is difficult to obtain. Monte Carlo techniques may be used to sample trajectories $\widetilde{\mathbf{x}}$, but hitting the formation is a rare event unless there is a mechanism to "guide" the trajectory into the formation. In this work, we determine the trajectory $\widetilde{\mathbf{x}}|\tau, \mathbf{x}_0$ conditioned on its hitting time. From the law of total expectation,

$$\Psi(\mathbf{x}) = \mathbb{E}_{\tau|\mathbf{x}_0}\left\{\mathbb{E}_{\widetilde{\mathbf{x}}|\tau,\widetilde{\mathbf{x}}_0}\left[\Psi(\widetilde{\mathbf{x}}(\tau))\exp\left(-\frac{1}{2\lambda}\int_0^{\tau} k(\widetilde{\mathbf{x}}(s))ds\right)\right]\right\} \tag{6.29}$$

$$= \mathbb{E}_{\tau|\mathbf{x}_0}\left\{\Psi(\mathbf{x}|\mathbf{x}_0,\tau)]\right\}. \tag{6.30}$$

In practice, we find that the inner distribution $\Psi(\mathbf{x}|\mathbf{x}_0,\tau)$ exhibits small tails for most τ and has high probability for just a small range of τ. Moreover, the range of τ with higher likelihood $\Psi(\mathbf{x}|\mathbf{x}_0,\tau)$ is that which appears to equally balance state and control costs. Therefore, we consider a discrete set $(\tau_1,\ldots,\tau_{N_\tau})$ of N_τ possible values for τ with non-informative, uniform prior probabilities. This implies that the distribution of the hitting times is implicitly encoded in the length (and the ensuing cost) of the path $\widetilde{\mathbf{x}}|\tau_i, \mathbf{x}_0$. Since $\Psi(\mathbf{x}(\tau_i)) = 1$ from (6.26), the solution (6.29) can be expanded as:

$$\Psi(\widetilde{\mathbf{x}}) = \frac{1}{N_\tau}\sum_{i=1}^{N_\tau}\mathbb{E}_{\widetilde{\mathbf{x}}|\widetilde{\mathbf{x}}_0,\tau_i}\left\{\Psi(\widetilde{\mathbf{x}}(\tau_i))\exp\left(-\frac{1}{2\lambda}\int_0^{\tau_i} k(\widetilde{\mathbf{x}}(s))ds\right)\right\} \tag{6.31}$$

$$= \frac{1}{N_\tau}\sum_{i=1}^{N_\tau}\mathbb{E}_{\widetilde{\mathbf{x}}|\widetilde{\mathbf{x}}_0,\tau_i}\left\{\exp\left(-\frac{1}{2\lambda}\int_0^{\tau_i} k(\widetilde{\mathbf{x}}(s))ds\right)\right\}. \tag{6.32}$$

By discretising the interval $[0,\tau_i]$ into N_i intervals of equal length Δt, $t_0 < t_1 < \ldots < t_{N_i} = \tau_i$, we can consider a sample of the discretised trajectory $\widetilde{\mathbf{x}}^N|\mathbf{x}_0,\tau_i = (\widetilde{\mathbf{x}}_1,\ldots,\widetilde{\mathbf{x}}_{N_i})$. Under this discretisation in time, the solution (6.32) with a right hand Riemann sum

approximation to the integral can be written as

$$\Psi(\widetilde{\mathbf{x}}) = \frac{1}{N_\tau} \lim_{\Delta t \to 0} \sum_{i=1}^{N_\tau} \int \mathrm{d}\widetilde{\mathbf{x}}^N P(\widetilde{\mathbf{x}}^N | \widetilde{\mathbf{x}}_0, \tau_i) \exp\left[-\frac{\Delta t}{2\lambda} \sum_{k=1}^{N_i} k(\widetilde{\mathbf{x}}_k) \right], \tag{6.33}$$

where $\mathrm{d}\widetilde{\mathbf{x}}^N = \prod_{k=1}^{N_i} \mathrm{d}\widetilde{\mathbf{x}}_k$ and where $P(\widetilde{\mathbf{x}}^N | \widetilde{\mathbf{x}}_0, \tau_i)$ is the probability of a discretised sample path, conditioned on the starting state $\widetilde{\mathbf{x}}_0$ and hitting time τ_i, given by

$$P(\widetilde{\mathbf{x}}^N | \widetilde{\mathbf{x}}_0, \tau_i) = \prod_{k=0}^{N_i - 1} p(\widetilde{\mathbf{x}}_{k+1} | \widetilde{\mathbf{x}}_k, \tau_i). \tag{6.34}$$

Since the uncontrolled process (6.28) is driven by Gaussian noise with zero mean and covariance $\Sigma = \Gamma\Gamma^\intercal$, the transition probabilities may be written as

$$\begin{aligned} p(\widetilde{\mathbf{x}}_{k+1} | \widetilde{\mathbf{x}}_k, \tau_i) \propto \exp\Big(& -\frac{1}{2} \left(\widetilde{\mathbf{x}}_{k+1} - \widetilde{\mathbf{x}}_k - f(\widetilde{\mathbf{x}}_k) \Delta t \right)^\intercal \\ & \times \left(\Delta t \lambda B R^{-1} B^\intercal \right)^{-1} \left(\widetilde{\mathbf{x}}_{k+1} - \widetilde{\mathbf{x}}_k - f(\widetilde{\mathbf{x}}_k) \Delta t \right) \Big) \end{aligned} \tag{6.35}$$

for $k < N_i - 1$, and $p(\widetilde{\mathbf{x}}_{k+1} | \widetilde{\mathbf{x}}_k, \tau_i) = \mathbb{1}_{h(\widetilde{\mathbf{x}}_{k+1}) = \boldsymbol{\mu}}(\widetilde{\mathbf{x}}_{k+1})$ for $k = N_i - 1$.

The path integral representation of $\Psi(\widetilde{\mathbf{x}})$ is obtained from equations (6.33)-(6.35), and can be written as an exponential of an "action" [111] $S(\widetilde{\mathbf{x}}^N | \widetilde{\mathbf{x}}_0, \tau_i)$ along the time-discretised sample trajectories $(\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_N)$:

$$\Psi(\widetilde{\mathbf{x}}) \propto \frac{1}{N_\tau} \lim_{\Delta t \to 0} \sum_{i=1}^{N\tau} \int d\widetilde{\mathbf{x}}^N \exp\left(-S(\widetilde{\mathbf{x}}^N | \widetilde{\mathbf{x}}_0, \tau_i) \right) \tag{6.36}$$

$$\begin{aligned} S(\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_N | \widetilde{\mathbf{x}}_0, \tau_i) = & \sum_{k=1}^{N_i} \frac{\Delta t}{2\lambda} k(\widetilde{\mathbf{x}}_k) + \sum_{k=0}^{N_i - 1} \frac{1}{2} \left(\widetilde{\mathbf{x}}_{k+1} - \widetilde{\mathbf{x}}_k - \Delta t f(\widetilde{\mathbf{x}}_k) \right)^\intercal \\ & \times \left(\lambda \Delta t B R^{-1} B^\intercal \right)^{-1} \left(\widetilde{\mathbf{x}}_{k+1} - \widetilde{\mathbf{x}}_k - \Delta t f(\widetilde{\mathbf{x}}_k) \right). \end{aligned} \tag{6.37}$$

Differentiating (6.36) with respect to $\widetilde{\mathbf{x}}_0$, we can obtain the optimal control (6.20)

as [271]

$$\begin{aligned}\mathbf{u}(\widetilde{\mathbf{x}}) &= \lim_{\Delta t\to 0} \lambda R^{-1}B^{\intercal}\partial_{\widetilde{\mathbf{x}}}\log\Psi \\ &= \lim_{\Delta t\to 0}\frac{1}{N_\tau}\sum_{i=1}^{N_\tau}\int d\widetilde{\mathbf{x}}^N P(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0,\tau_i)\mathbf{u}_L(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0,\tau_i) \qquad (6.38)\\ &= \lim_{\Delta t\to 0}\frac{1}{N_\tau}\sum_{i=1}^{N_\tau}\mathbb{E}_{P(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0,\tau_i)}\left\{\mathbf{u}_L(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0,\tau_i)\right\} \qquad (6.39)\end{aligned}$$

where $\lim_{\Delta t\to 0} P(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0,\tau_i) = P(\widetilde{\mathbf{x}}|\widetilde{\mathbf{x}}_0,\tau_i)$ is the probability of an *optimal* trajectory conditioned to hit the formation at time τ_i,

$$P(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0,\tau_i)\propto e^{-S(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0,\tau_i)}, \qquad (6.40)$$

which weights the local controls $\mathbf{u}_L(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0,\tau_i)$ in (6.38), defined by

$$\mathbf{u}_L(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0,\tau_i) = R^{-1}B^T\left(BR^{-1}B^T\right)^{-1}\left(\frac{\widetilde{\mathbf{x}}_1-\widetilde{\mathbf{x}}_0}{\Delta t}-f(\widetilde{\mathbf{x}}_0)\right). \qquad (6.41)$$

Although the resulting control law is stationary, the state space is too large for it to be computed offline. Because of this, after computing $\mathbf{u}(\mathbf{x}) = \mathbf{u}(\widetilde{\mathbf{x}}_0)$, each agent executes only the first increment of that control, at which point the optimal control is recomputed. Then (6.39) is

$$\begin{aligned}\mathbf{u}(\mathbf{x}) &= R^{-1}B^T\left(BR^{-1}B^T\right)^{-1}\left(\frac{\mathbb{E}_{P(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0)}\{\widetilde{\mathbf{x}}_1\}-\mathbf{x}}{\Delta t}-f(\mathbf{x})\right)\\ &= R^{-1}B^T\left(BR^{-1}B^T\right)^{-1}\left(\frac{\mathbb{E}_\tau\mathbb{E}_{P(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0,\tau)}\{\widetilde{\mathbf{x}}_1\}-\mathbf{x}}{\Delta t}-f(\mathbf{x})\right). \qquad (6.42)\end{aligned}$$

In other words, the control (6.42) applied by an agent in state $\mathbf{x}$ is constructed from a realisation of the unknown or random dynamics of the system that maximizes the probability of the trajectory that starts from $\mathbf{x}$ and evolves until hitting the formation. This probability is weighted by the cost accumulated along the path.

6.3.2 Discounted cost infinite-horizon control

For the discounted infinite-horizon costs, we will develop the derivation separately while showing the relation to the CUF problem. We begin by writing the cost functional (6.17) in terms of a finite-horizon cost functional and an error term ε:

$$\begin{aligned} J(\mathbf{x},t) &= \min_{\mathbf{u}} \mathbb{E}\left\{ \int_0^T \frac{e^{-\beta s}}{2} \left(k(\mathbf{x}) + \mathbf{u}^\mathsf{T} R \mathbf{u}\right) ds \right\} + \varepsilon \\ &= \min_{\mathbf{u}} \mathbb{E}\left\{ \int_0^T \frac{1}{2} \left((h(\mathbf{x}) - \boldsymbol{\mu})^\mathsf{T} Q_t (h(\mathbf{x}) - \boldsymbol{\mu}) + \mathbf{u}^\mathsf{T} R_t \mathbf{u}\right) ds \right\} + \varepsilon, \end{aligned} \tag{6.43}$$

where it is assumed that $T > \overline{T}$ is sufficiently large so that ε may be neglected, and where $Q_t = e^{-\beta t} Q$ and $R_t = e^{-\beta t} R$. The feedback control $\mathbf{u}$ that minimizes (6.43) will be used in a receding horizon manner over the horizon $[0,T]$, i.e., $\mathbf{u}(\mathbf{x}) = \mathbf{u}(\mathbf{x}(0))$.

The (stochastic) HJB equation for the system kinematics (6.12) and cost functional (6.43), assuming $\varepsilon = 0$, is

$$\begin{aligned} 0 = \partial_t J + \min_{\mathbf{u}} \Bigg(& (f + B\mathbf{u})^\mathsf{T} \partial_{\mathbf{x}} J + \frac{1}{2} \mathrm{Tr}\left(\Sigma \partial_{\mathbf{x}}^2 J\right) \\ & + \frac{e^{-\beta t}}{2} k(\mathbf{x}(t)) + \frac{1}{2} \mathbf{u}(\mathbf{x})^\mathsf{T} R_t \mathbf{u}(\mathbf{x}) \Bigg). \end{aligned} \tag{6.44}$$

Substituting the optimal control $\mathbf{u}(\mathbf{x},t) = -R_t^{-1} B^\mathsf{T} \partial_{\mathbf{x}} J(\mathbf{x},t)$ and applying the transformation $J(\mathbf{x},t) = -\lambda \log \Psi(\mathbf{x},t)$, we obtain

$$\begin{aligned} & \frac{1}{\Psi} \partial_t \Psi = e^{-\beta t} \frac{k(\mathbf{x})}{2\lambda} - \frac{f^\mathsf{T}}{\Psi} \partial_{\mathbf{x}} \Psi - \frac{1}{2} \frac{1}{\Psi} \mathrm{Tr}\left(\Sigma \partial_{\mathbf{x}}^2 \Psi\right) \\ & - \frac{1}{2} \frac{\lambda}{\Psi^2} \left(\partial_{\mathbf{x}} \Psi\right)^\mathsf{T} B R_t^{-1} B^\mathsf{T} \partial_{\mathbf{x}} \Psi + \frac{1}{2} \frac{1}{\Psi^2} \left(\partial_{\mathbf{x}} \Psi\right)^\mathsf{T} \Sigma_t \partial_{\mathbf{x}} \Psi. \end{aligned} \tag{6.45}$$

In order for the nonlinear terms to cancel as in (6.22), we use the assumption (6.23),

but with time dependence due to the discounting factor included:

$$\Sigma_t = e^{\beta t}\Gamma\Gamma^\intercal = \lambda B R_t^{-1} B^\intercal. \tag{6.46}$$

Note that in a feedback setting, the control $\mathbf{u}(\mathbf{x})$ is based on the state $\mathbf{x}(0)$, and Σ_0 reduces to $\lambda B R^{-1} B^\intercal$. The remaining PDE,

$$\partial_t \Psi = \left(e^{-\beta t} \frac{k(\mathbf{x})}{2\lambda} - f^\intercal \partial_{\mathbf{x}} - \frac{1}{2} \mathrm{Tr}\left(\Sigma_t \partial_{\mathbf{x}}^2\right) \right) \Psi, \tag{6.47}$$

is linear and has solution

$$\Psi(\widetilde{\mathbf{x}}_0, 0) = \mathbb{E}_{\widetilde{\mathbf{x}}|\widetilde{\mathbf{x}}_0} \left\{ \exp\left(-\frac{1}{2\lambda} \int_0^T e^{-\beta s} k(\mathbf{x}(s)) ds \right) \right\}. \tag{6.48}$$

The reader will note that this solution also appears in the case of the CUF problem in (6.32), but with a few changes. First, only one horizon T is considered, and so the summation over τ_i is removed. Next, the discounting factor β is included inside the expectation. Finally, the uncontrolled process $\widetilde{\mathbf{x}}$ is (6.28) but with an increasing noise intensity due to (6.46):

$$\mathrm{d}\widetilde{\mathbf{x}}(t) = f(\widetilde{\mathbf{x}}(t))\mathrm{d}t + e^{\frac{\beta t}{2}} \Gamma \mathrm{d}\mathbf{w}. \tag{6.49}$$

The path integral representation of $\Psi(\widetilde{\mathbf{x}}_0, 0)$ corresponding to the discounted infinite-

horizon problem is obtained in the same manner as (6.36)-(6.37), and is

$$\Psi(\widetilde{\mathbf{x}}_0,0) \propto \lim_{\Delta t\to 0}\int \mathrm{d}\widetilde{\mathbf{x}}^N \exp\left(-S(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0)\right) \tag{6.50}$$

$$\begin{aligned} S(\widetilde{\mathbf{x}}_1,\ldots,\widetilde{\mathbf{x}}_N|\widetilde{\mathbf{x}}_0) &= \sum_{k=1}^{N}\frac{e^{-\beta t_k}\Delta t}{2\lambda}k(\widetilde{\mathbf{x}}_k) \\ &+ \sum_{k=0}^{N-1}\frac{1}{2}\left(\widetilde{\mathbf{x}}_{k+1}-\widetilde{\mathbf{x}}_k-\Delta t f(\widetilde{\mathbf{x}}_k)\right)^{\intercal} \\ &\times\left(\lambda\Delta t B R^{-1}B^{\intercal}e^{\beta t_k}\right)^{-1}\left(\widetilde{\mathbf{x}}_{k+1}-\widetilde{\mathbf{x}}_k-\Delta t f(\widetilde{\mathbf{x}}_k)\right). \end{aligned} \tag{6.51}$$

Along similar lines, the control may be computed as (cf. (6.42))

$$\mathbf{u}(\mathbf{x}) = R^{-1}B^T\left(BR^{-1}B^T\right)^{-1}\left(\frac{\mathbb{E}_{P(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0)}\{\widetilde{\mathbf{x}}_1\}-\mathbf{x}}{\Delta t}-f(\mathbf{x})\right) \tag{6.52}$$

where the path probability is

$$P(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0)\propto e^{-S(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0)}. \tag{6.53}$$

Both for the CUF problem and the infinite horizon problem, the matrix coefficients multiplying (6.42) and (6.52) may be dropped since control is only affecting the noisy states[4]. One may compute the optimal control (6.42) (resp. (6.52)) once the path probability $P(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0,\tau_i)$ (6.40) (resp. (6.53)) has been computed, a nontrivial task to be discussed in the following section.

6.4 Computing the Control with Kalman Smoothers

In this section we present our approach to compute the control in (6.42) and (6.52). Although Monte Carlo techniques can be used to generate samples of the maximally-

[4]The reader may examine the form of $B\mathbf{u}$, taking into account that the components of the final multiplicative factor in (6.42) or (6.52) are zero for uncontrolled states.

likely trajectory $P(\widetilde{\mathbf{x}}_1,\ldots,\widetilde{\mathbf{x}}_N|\widetilde{\mathbf{x}}_0)$, we find them to be slow in practice due to the high dimension of this problem ($\widetilde{\mathbf{x}}^N \in \mathbb{R}^{4NM}$). Moreover, in the case of CUF, when sampling a trajectory $\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0,\tau_i$, the trajectory must be conditioned to hit the formation at τ_i. Finally, it is not necessary to sample the entire distribution $P(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0)$ since only the estimate $\widehat{\mathbf{x}}_1 \equiv \mathbb{E}_{P(\widetilde{\mathbf{x}}^N|\widetilde{\mathbf{x}}_0)}\{\widetilde{\mathbf{x}}_1\}$ is needed.

Therefore, in this work, we treat the temporal discretisation of the optimal trajectory $\widetilde{\mathbf{x}}^N$ as the hidden state of a stochastic process, where appropriately-chosen measurements of this hidden state are related to the system goal $\boldsymbol{\mu}$ (6.16). The optimal control can then be computed from the optimal estimate $\widehat{\mathbf{x}}_1$ given the process and measurements over a fixed interval $t_1,\ldots,\tau_i$ for the CUF problem and $t_1,\ldots,T$ for the infinite-horizon problem. We define the following nonlinear smoothing problem.

Nonlinear Smoothing Problem:

Given measurements $\mathbf{y}_k = \mathbf{y}(t_k)$ for $t_k = t_1,\ldots,t_N = \tau_i$ (or T), where $t_{k+1} - t_k = \Delta t$, compute the estimate $\widehat{\mathbf{x}}_{1:N}$ of the hidden state $\widetilde{\mathbf{x}}_{1:N}$ from the nonlinear hidden state-space model:

$$\widetilde{\mathbf{x}}_{k+1} = \widetilde{\mathbf{x}}_k + \Delta t\, f(\widetilde{\mathbf{x}}_k) + \boldsymbol{\varepsilon}_k \tag{6.54}$$

$$\mathbf{y}_k = h(\widetilde{\mathbf{x}}_k) + \boldsymbol{\eta}_k, \tag{6.55}$$

where $f(\cdot)$ and $h(\cdot)$ are as in Section 6.2, and $\boldsymbol{\varepsilon}_k$ and $\boldsymbol{\eta}_k$ are independent multivariate

Gaussian random variables with zero mean and with covariances chosen as follows:

Control until formation:

$$\mathbb{E}(\varepsilon_k \varepsilon_k^\intercal) = \lambda \Delta t B R^{-1} B^\intercal \tag{6.56}$$

$$\mathbb{E}(\eta_k \eta_k^\intercal) = \begin{cases} \frac{\lambda}{\Delta t} Q^{-1} & k = 1, \ldots, N-1 \\ 0 & k = N \end{cases}. \tag{6.57}$$

Discounted infinite horizon:

$$\mathbb{E}(\varepsilon_k \varepsilon_k^\intercal) = \lambda \Delta t B R^{-1} B^\intercal e^{\beta t_k} \tag{6.58}$$

$$\mathbb{E}(\eta_k \eta_k^\intercal) = \frac{\lambda}{\Delta t} Q^{-1} e^{\beta t_k}. \tag{6.59}$$

The smoothing is initialized from $\widetilde{\mathbf{x}}_0 = \mathbf{x}$, the current state of the system as viewed by the agent. Measurements $\mathbf{y}_k$ are always exactly $\mathbf{y}_k = \boldsymbol{\mu}$. □

Note that the measurement and process covariances have the same structure for both cost function types, but the noise intensity increases during the smoothing horizon in the discounted infinite horizon case. As the covariance becomes large, the measurements and state predictions contain so little information that the change in the estimated hidden state (the trajectory) is negligible, leading to a stationary control policy.

To show the relation between the nonlinear smoothing problem and the stochastic optimal control problem, we write the probability of a hidden state sequence $\widetilde{\mathbf{x}}^N$ given measurements $\mathbf{y}_k$ and the initial state $\widetilde{\mathbf{x}}_0$, which is [108]

$$P\left(\widetilde{\mathbf{x}}^N | \widetilde{\mathbf{x}}_0, \mathbf{y}_1, \ldots, \mathbf{y}_N\right) \propto \prod_{k=1}^{N} p(\mathbf{y}_k | \widetilde{\mathbf{x}}_k) p(\widetilde{\mathbf{x}}_k | \widetilde{\mathbf{x}}_{k-1}), \tag{6.60}$$

where, in the case of the discounted infinite-horizon cost function[5],

$$
\begin{aligned}
p(\mathbf{y}_k|\widetilde{\mathbf{x}}_k) \equiv p(\boldsymbol{\mu}_k|\widetilde{\mathbf{x}}_k) &= N\left(h(\widetilde{\mathbf{x}}_k), \eta_k \eta_k^{\intercal}\right) \\
&\propto \exp\left\{-\frac{\Delta t}{2\lambda}(h(\widetilde{\mathbf{x}}_k)-\boldsymbol{\mu})^{\intercal} Q e^{-\beta t}(h(\widetilde{\mathbf{x}}_k)-\boldsymbol{\mu})\right\} \qquad (6.61) \\
p(\widetilde{\mathbf{x}}_k|\widetilde{\mathbf{x}}_{k-1}) &= N\left(\widetilde{\mathbf{x}}_{k-1}+\Delta t f(\widetilde{\mathbf{x}}_{k-1}), \Delta t \Sigma\right) \\
&\propto \exp\left\{-\frac{1}{2}\left(\widetilde{\mathbf{x}}_k-\widetilde{\mathbf{x}}_{k-1}-\Delta t f(\widetilde{\mathbf{x}}_{k-1})\right)^{\intercal}\right. \\
&\qquad \left.\times\left(\lambda \Delta t B R^{-1} B^{\intercal} e^{\beta t}\right)^{-1}\left(\widetilde{\mathbf{x}}_k-\widetilde{\mathbf{x}}_{k-1}-\Delta t f(\widetilde{\mathbf{x}}_{k-1})\right)\right\}. \qquad (6.62)
\end{aligned}
$$

Comparing the right hand sides of (6.61)-(6.62) with (6.50)-(6.51) (resp. (6.36)-(6.37)), it can be seen that they are identical to those in the stochastic optimal control problem.

Since the optimal control (6.42) is based on the probability of a full trajectory of fixed length and values $\boldsymbol{\mu}$ are available in advance, the expected value of the trajectory originating from state $\widehat{\mathbf{x}}_0$ conditioned to hit the formation at time τ_i or T, that is, the hidden states $\widehat{\mathbf{x}}_k$, $k = 1,\ldots,N$, can be found by filtering and then smoothing the process given the values $\boldsymbol{\mu}_k$ using a nonlinear fixed-interval Kalman smoother. Such an algorithm assumes that the increments given by (6.61) and (6.62) are Gaussian[6] to some extent, but the algorithm is sufficiently fast to be applied in *real-time* by each unicycle in a potentially large group with an even larger state space, motivating its use in this work.

In the case of the first-hitting time problem, the estimated trajectory is first computed for each τ_i, at which point the expectation over τ_i may be computed. This results

[5] Set $\beta = 0$ for the CUF case

[6] See [154] for a discussion on the Gaussianity of these increments.

in an average of the controls $u_L(\widetilde{\mathbf{x}}|\widetilde{\mathbf{x}}_0, \tau_i)$ to be applied. In other words, each agent estimates both the optimal system trajectory (from its perspective) given the time the formation will hit *and* the hitting time of the formation. Hitting the target sooner would save on state costs, but may cause an increase in control costs, and vice versa.

When the smoothing is complete and agents have applied their computed control, each agent must then observe the actual states of its neighbours so that the next iteration begins with the correct initial condition. We provide a pseudo code for our computations in Algorithm 1 for the CUF case and in Algorithm 2 for the infinite horizon case. For clarify, we also provide a flowchart in Figure 6.2.

Algorithm 1 Formation control algorithm applied by each agent: control to formation case

$\mathbf{x}(t) \leftarrow$ measured state of system from agent's viewpoint
$\boldsymbol{\mu} \leftarrow$ nominal distances
while $\mathbf{x}(t) \notin \mathcal{F}$ **do**
  for $i = 1, \ldots, N_\tau$ **do**
    $\mathbb{E}\{\widetilde{\mathbf{x}}_1\}, \ldots, \mathbb{E}\{\widetilde{\mathbf{x}}_N\} \leftarrow$ KalmanSmoother(initial state $= \mathbf{x}_0 = \mathbf{x}$, horizon $= [0, \tau_i]$, predictions (6.28) or (6.49), measurements $= \boldsymbol{\mu}$)
    $\mathbb{E}\left\{\mathbf{u}_L(\widetilde{\mathbf{x}}^N|\mathbf{x}_0, \tau_i)\right\} \leftarrow (\mathbb{E}(\widetilde{\mathbf{x}}_1) - \mathbf{x})/\Delta t - f(\mathbf{x})$ ▷ from (6.41)
  end for
  $\mathbf{u}(\mathbf{x}) = \frac{1}{N_\tau}\sum_{i=1}^{N_\tau} \mathbb{E}\left\{\mathbf{u}_L(\widetilde{\mathbf{x}}^N|\mathbf{x}_0, \tau_i)\right\}$
  Apply computed corrective control $\mathbf{u}(\mathbf{x})$ ▷ using (6.12)
  $\mathbf{x}(t) \leftarrow$ measured state of system from agent's viewpoint
end while

In practice, the controller/smoother must be capable of efficiently smoothing over the horizon $[t_0, \tau_i]$ (or $[t_0, T]$). The computational complexity [218] of the smoother used in this work [219] roughly scales with the number of neighbours as M^3. As such, for implementation on mobile robotic controllers, the computational requirements may

Algorithm 2 Formation control algorithm applied by each agent: infinite horizon case

$\mathbf{x}(t) \leftarrow$ measured state of system from agent's viewpoint
$\boldsymbol{\mu} \leftarrow$ nominal distances
loop
  $\mathbb{E}\{\widetilde{\mathbf{x}}_1\}, \ldots, \mathbb{E}\{\widetilde{\mathbf{x}}_N\} \leftarrow$ KalmanSmoother(initial state $= \mathbf{x}_0 = \mathbf{x}$,
    horizon $= [0, T]$,
    predictions from (6.28) or (6.49),
    measurements $= \boldsymbol{\mu}$)
  $\mathbf{u}(\mathbf{x}) = \mathbb{E}\left\{\mathbf{u}_L(\widetilde{\mathbf{x}}^N|\mathbf{x}_0)\right\} \leftarrow (\mathbb{E}(\widetilde{\mathbf{x}}_1) - \mathbf{x})/\Delta t - f(\mathbf{x})$ ▷ from (6.52)
  Apply computed corrective control $\mathbf{u}(\mathbf{x})$ ▷ using (6.12)
  $\mathbf{x}(t) \leftarrow$ measured state of system from agent's viewpoint
end loop

prescribe a two-level approach – a high-level nonlinear fixed-interval Kalman smoothing algorithm that computes reference inputs for lower-level robot control loops. However, various implementations of the smoothing algorithm can streamline the computation. For instance, a sigma point-based approximation to the two-filter smoothing equations [279] could produce reference control values using only twice the computing power of an unscented filtering algorithm. Moreover, through linearisation of the measurement model (6.55) or of both the process and measurement models (6.54)-(6.55), the control could be computed using extended Kalman smoother or the linear unscented Kalman smoother, respectively. However, through full linearisation, we would effectively only be solving an LQG control problem, as mentioned in Section 6.1.

6.5 Results

In this section we compare the control computed by our Kalman smoothing algorithm against that computed by a numerical stochastic optimal control method [137], and

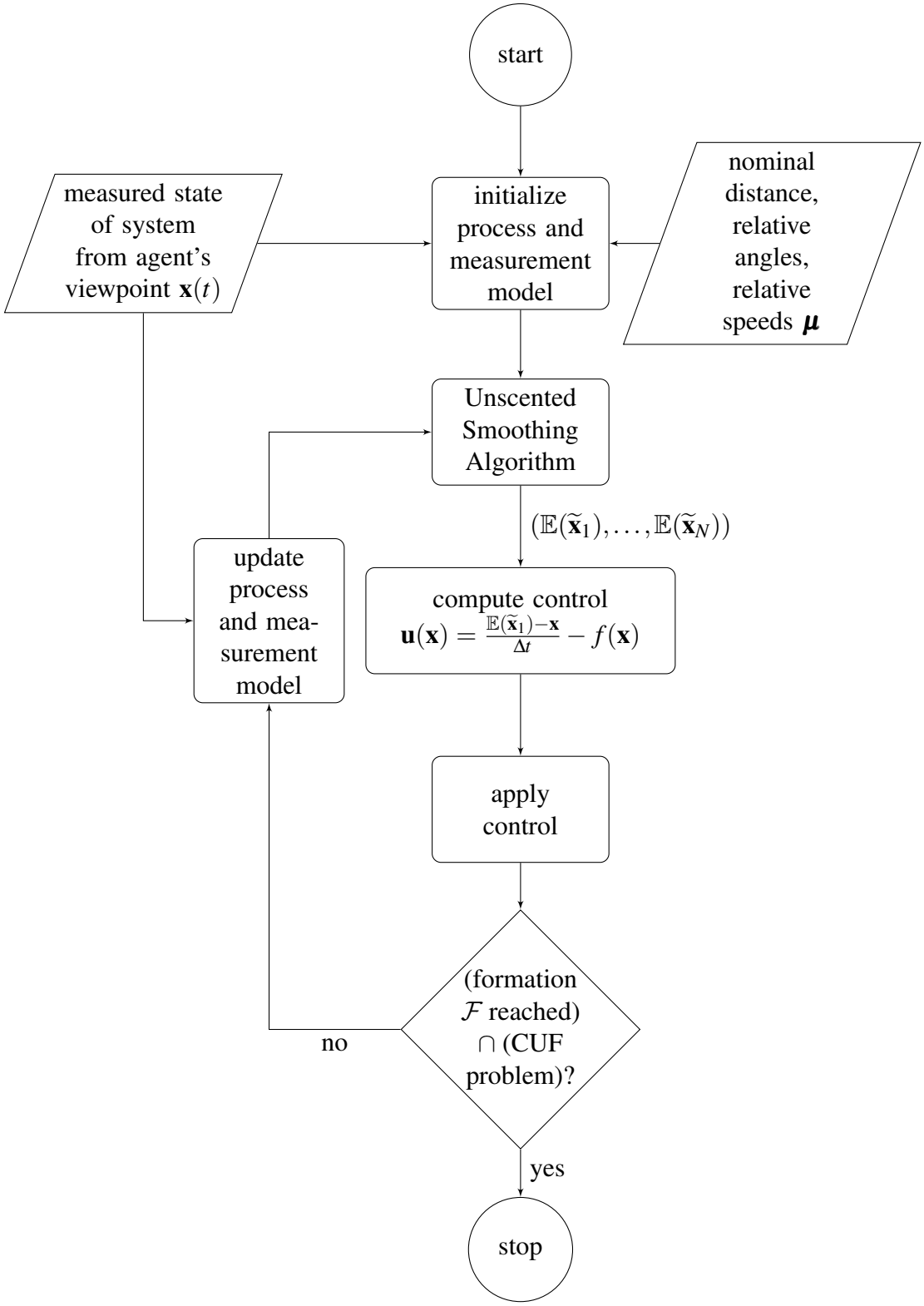


Figure 6.2: Flowchart for formation control algorithm

then we provide simulations in which five agents achieve formations with equal and aligned velocities. The system and algorithm parameters were chosen as $\lambda = 100$, $\sigma_\theta = \sigma_{\theta,m} = 0.1$ (see (6.24) for R), $Q = 100I$, and $\Delta t = 0.1$, with all units relative to meters and seconds. In the case of CUF, $N_\tau = 10$, $\tau = 1,\ldots,10$, $\varepsilon_r = \varepsilon_v = 0.1$, $\varepsilon_\theta = 10^\circ$. For the infinite-horizon case, $\beta = 0.5$ and $T = 10$. With these parameters, $\exp(-\beta T) \approx 0.007$, and so the cost function (6.43) used for the algorithm is a good approximation to a discounted infinite-horizon cost function (6.17). In each case, the control was computed using a Discrete-time Unscented Kalman Rauch-Tung-Striebel Smoother [218, 219].

We first apply the Markov chain approximation method [137] to compute the feedback control from the HJB equation for cost function (6.17) so that we may compare the computation time and computed controls with our proposed method. In order to reduce the dimension of the state space for the Markov chain approximation method, which requires discretisation of the state space, we consider the problem of just two agents, fix $v_1 = v_2 = 1$ [m/s], and aim for a nominal distance between agents of $\delta_1 = 5$ [m]. Then the controls $\omega(\Delta x_1, \Delta y_1, \theta, \theta_1)$ and $\omega_1(\Delta x_1, \Delta y_1, \theta, \theta_1)$ may be computed, although this is computationally intensive due to the discretisation of the state space ($\mathbf{x} \in \mathbb{R}^4$) and the control variable $\mathbf{u} \in \mathbb{R}^2$, leading to 25,600 grid cells and 225 control possibilities for the chosen discretisation. The left column of Figure 6.3 shows the optimal controls as computed by the Markov chain approximation method, which, for the chosen parameters, required 3.7 hours of computation time in Matlab® on an Intel i7 at 2.7 GHz with 8

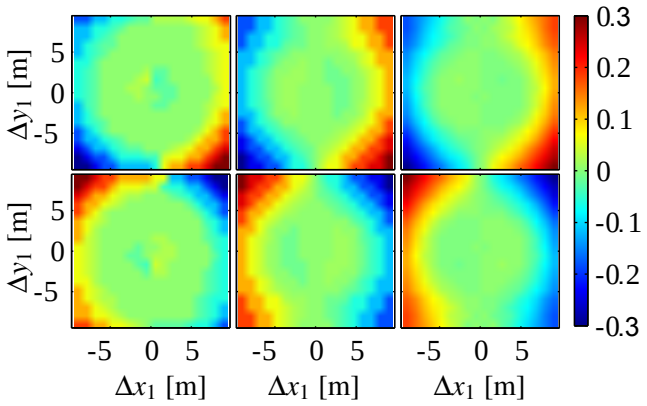


Figure 6.3: Optimal feedback control $\mathbf{u}(\Delta x_1, \Delta y_1, \theta_1, \theta_2)$ based on a discounted infinite horizon for two agents with fixed speed $v = 1$, evaluated at $\theta_1 = \theta_2 = 0$. The top row is the control ω and the bottom is ω_1 (i.e., the control assumed for a neighbour $m = 1$). (Left column) Control based on Markov chain approximation with Δx step size $= \Delta y$ step size $= 1.05$, θ_1 step size $= \theta_2$ step size $= .078$, and control $\mathbf{u}$ step size 0.85 for both ω and ω_1. (Right column) Control evaluated using a Kalman smoother on the same grid, but without control discretisation. (Center column) To compare, Kalman smoother control down-sampled to the same control step sizes 0.85 for ω and ω_1 that are used in the left column's control computation.

GB RAM (0.51 s per grid cell visible in Figure 6.3). To compare, the Kalman smoother algorithm described in the previous section was applied to the same grid locations in state space, but without the control discretisation. The Kalman smoother-based method required only 37.26 [s] of computation time, an average of 0.09 s per grid cell, although it is not actually necessary to compute the control for the entire state space grid.

Figure 6.3 serves two purposes. First, it is seen that the controls computed by our method are similar to those from the Markov chain approximation method, although with such an aggressive discretisation required by the latter, there are no guarantees that the control solutions should be identical. Moreover, recalling that the displayed controls are additive corrections to the non-optimal feedback controls, this figure allows

us to examine the regions in state space where the non-optimal feedback control is most deficient in terms of optimality and robustness to uncertainty. For example, in Figure 6.3 shows that little correction control is needed in regions $r \lesssim 5$, (i.e., $\mathbf{u} \approx 0$). However, as agents become further separated, the reference feedback controls are not sufficiently strong. Of course, the described corrections are only for the case $\theta_1 = \theta_2 = 0$ and with fixed speed, and we do not attempt in this paper to analytically improve the reference control nor its underlying artificial potential function. However, we envision our method being useful for control designers to analyse a “slice" or marginalisation of their feedback control function without having to solve the HJB PDE in the entire state space.

Next, the Kalman smoothing method is employed so that five agents achieve the formation of a regular pentagon, where each agent is individually estimating the hidden optimal trajectory based on the relative kinematics of all of its neighbours. The agents observe all others, but, as described in Section 6.2, only the inter-agent connections *known* by an agent are used when computing the control. The instantaneous state cost (6.14) penalizes the mean squared distance from the unicycle to all of its $M = 4$ neighbours in excess of the side length of the pentagon (5 [m]) or the diagonal of the pentagon, depending on the relative configuration of the pentagon encoded in δ_m, $m = 1, \ldots, 4$. The system parameters are the same as the previous example, with the addition of $\sigma_v = \sigma_{v,m} = 0.1$ (see (6.24) for R).

For the case of CUF, we define the formation, i.e., stopping condition, as

$$\begin{aligned}\mathcal{F} = \big\{ \mathbf{x} : & |r_m - \delta_m| \leq \varepsilon_r, \\ & |\theta - \theta_m| \leq \varepsilon_\theta, \\ & |v - v_m| \leq \varepsilon_v, \ \ m = 1, \ldots, M \big\}. \end{aligned} \tag{6.63}$$

based on tolerances ε_r, ε_θ, and ε_v. Figure 6.4 shows the trajectories of all agents, while the inter-agent distances and agents' angles and speeds can be seen in Figure 6.5. The actual stopping time was $\tau = 16.1$ [s], and the agents did not collide. In the last second of the simulation, a minor deviation in the agents' heading angles and speeds was seen. This was due to a final correction in agents' relative distances that was needed after having converged in heading angle and speed. Without the addition of the optimal controls, the collision-avoiding controls acting alone did not lead to a formation (6.63) within the first 60 [s] of the simulation (Figure 6.5). Although the non-optimal control aligned heading angles and, to some extent, speeds, they led to oscillatory trajectories, which is not uncommon in the artificial potential function approach.

The infinite-horizon case may be seen in Figure 6.6, and its corresponding cost is in Figure 6.7. The agents achieve and maintain a pentagon using the computed optimal controls, while a simulation using only the reference feedback controls again leads to oscillations, as seen in Figure 6.7.

As discussed in Section 6.1, the implementation of the Kalman smoothing algorithm is in some ways similar to receding horizon control/model predictive control. To compare our results against this strategy, we next consider an Euler discretisation of the

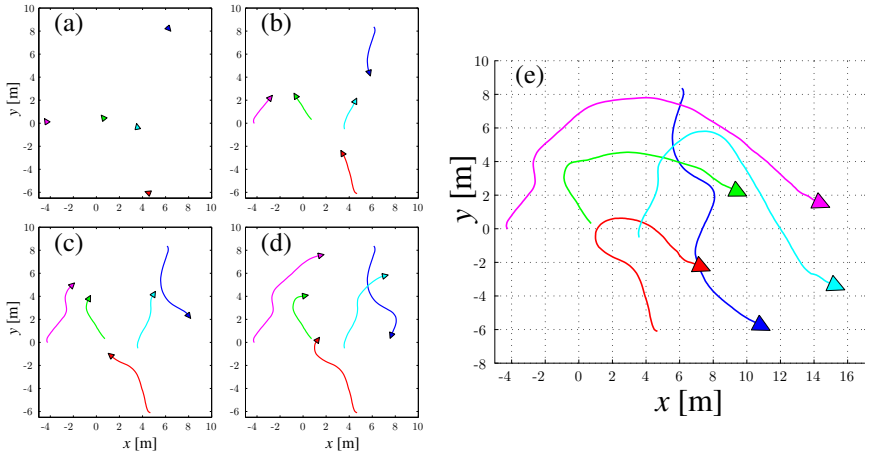


Figure 6.4: Five agents, starting from random initial positions and a common speed $v = 2.5$ [m/s], achieve a regular pentagon formation by an individually-optimal choice of acceleration and turning rate, without any active communication. The frames (a) through (e) correspond to the times $t = 0$ [s], $t = 2$ [s], $t = 4$ [s], $t = 6$ [s], and $t = 16.4$ [s]. An example of collision avoidance between the two upper-left agents is seen in frames (b) and (c).

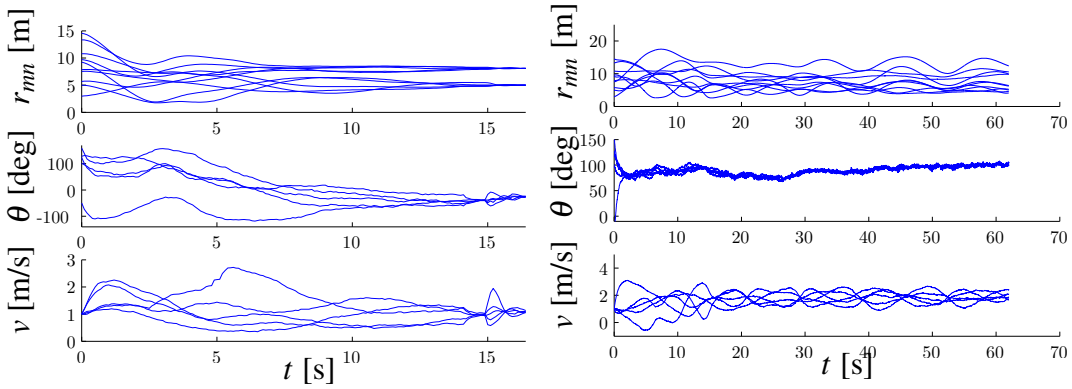


Figure 6.5: Inter-agent distances r_{mn}, agent heading angles θ, and agent speeds v as a function of time using the (left) stochastic optimal control (CUF) and the (right) deterministic feedback control.

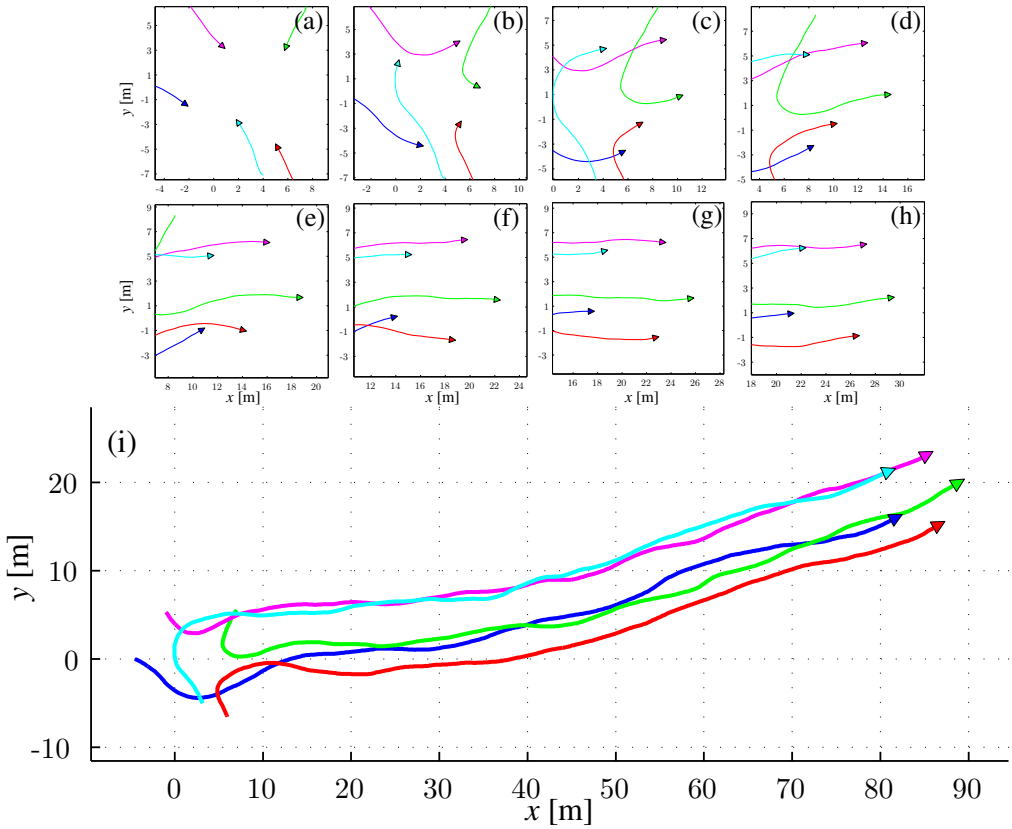


Figure 6.6: Five agents, starting from random initial positions and a common speed $v = 1$ [m/s], achieve a regular pentagon formation by an individually-optimal choice of acceleration and turning rate, without any active communication. Frames from (a) to (h) are 2.5 [s] to 20 [s], incrementing by 2.5 [s]. Frame (i) shows the end of the simulated trajectories at 60.6 [s].

system (6.12) with noise intensity $\Gamma = 0$. At each time step, each agent must compute the controls $(\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{N-1})$ that minimize the total accumulated cost (6.43) over the horizon, at which point the first control $\mathbf{u}_0$ is executed, and then the process repeats. Using the same time step (0.1 [s]) that was employed with the Kalman smoothers was computationally prohibitive, and so we chose $\Delta t = 0.5$ [s], resulting in 200 decision variables. We computed the control using IPOPT [278], which required approximately 30 minutes for each time step. Figure 6.8(a) shows that the agents under MPC control

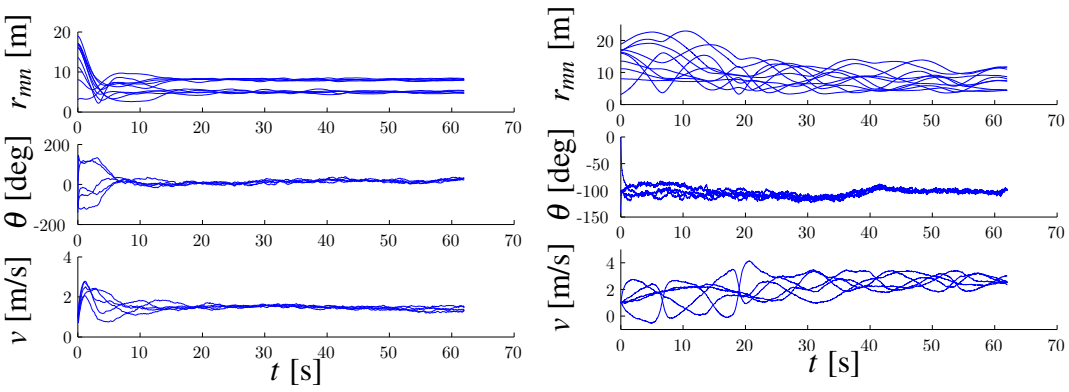


Figure 6.7: Inter-agent distances r_{mn}, agent heading angles θ, and agent speeds v as a function of time using the (left) stochastic optimal control (infinite-horizon) and the (right) deterministic feedback control

converge to a pentagon that moves in a different direction than the Kalman smoothing-based pentagon, but the direction of the agents' motion was not part of the optimization problem. The difference between the actual inter-agent distances and the nominal distances for the MPC-based trajectories are close to that from our Kalman smoothing methods (Figure 6.8(b)), but the latter approach requires considerably less computational time (approximately 1.7 [s] to compute five agents' controls).

Finally, we show how the choice of measurements $\boldsymbol{\mu}$ supplied to the smoothing algorithm can allow for dynamic morphing between formations. At the end of the simulation in Figure 6.6, we update the nominal inter-agent distances so that agents form a line, as seen in Figure 6.9.

6.6 Discussion

This work considers the problem of unicycle formation control in a distributed optimal feedback control setting. Since this gives rise to a system with huge state space,

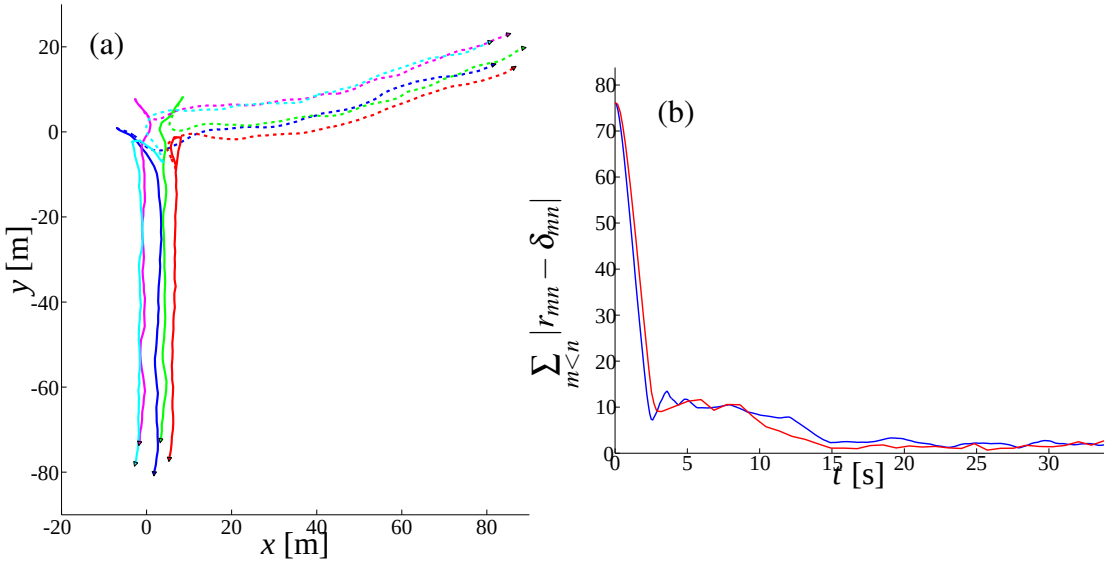


Figure 6.8: Comparison against MPC-computed control for the same problem. (a) The dashed trajectories are those from Figure 6.6 computed using Kalman smoothing algorithms, while the solid trajectories were computed using IPOPT [278]. (b) The total error between the instantaneous formation distances and the nominal distances for the MPC control (red) and the Kalman smoother control (blue).

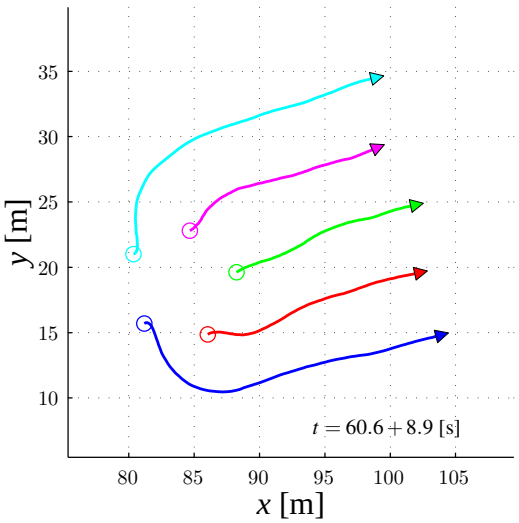


Figure 6.9: Following the simulation in Figure 6.6, the set of nominal distances between agent m and n was modified to $\delta_{mn} = 5|m-n|$, and the reference control was dropped. The agents are seen forming a line.

we exploit the stochasticity inherent in distributed multi-agent control problems in order to apply a path integral method. By relating the resulting estimation problem to Kalman smoothing algorithms, each agent can compute its optimal control using a nonlinear Kalman smoother, greatly reducing the computational complexity associated with multi-robot optimal control problems. The measurement and process noise of the smoothing problem are created using the structure of the cost function and stochastic kinematics. Since agents share a common prior probability distribution describing the future trajectories of their neighbours, the formation can be attained without any communication among agents aside from instantaneous observations of neighbours.

The typical Lyapunov function approach to problems of this type offers several benefits, including transparency and analyticity, that are often lost in the numerical solution of the Hamilton-Jacobi-Bellman equation, but the latter is also advantageous or in some instances preferable. Our approach aims to balance the attractiveness of an analytic feedback control for deterministic systems with the robustness to stochasticity and optimality provided by the Hamilton-Jacobi-Bellman equation. Therefore, the optimal turning rate and acceleration controls affect the system alongside a non-optimal reference feedback control law based on an artificial potential function.

Not only do the optimal correction terms improve the performance of the reference control, leading the agents into formation faster and optimally, but the existence of an underlying non-optimal control complements the optimal enhancement controls. For instance, since the reference controls help to move the agents toward a formation, a

Kalman smoother measurement (recall that a measurement in our work is the nominal distance) does not constitute a rare event in the measurement probability model, and, consequently, the reference controls reduce numerical issues involved in Kalman smoothing algorithms. Moreover, without the reference controls, collision avoidance would require complex state constraints to be added to the Kalman smoothing algorithm.

Since the Kalman smoothing-based control is the additional input required alongside the non-optimal control in order to achieve optimality and robustness, we can envision in a future line of research using these algorithms to test feedback control laws for multi-robot systems against various scenarios in order to derive analytical improvements.

Chapter 7

Self-Triggered p-moment Stability for Continuous Stochastic State-Feedback Controlled Systems

This chapter is a preprint to the paper

- Anderson, R. P., Milutinović, D., and Dimarogonas, D.V., "Self-Triggered p-moment Stability for Continuous Stochastic State-Feedback Controlled Systems", submitted to the *IEEE Transactions on Automatic Control*.

7.1 Introduction

The implementation of nonlinear feedback controllers on digital computer platforms necessitates a choice of time points at which the controller should be updated. In traditional setups, these updates are scheduled periodically with conservative, short time periods that under all circumstances do not compromise the stability of closed loops. However, a longer period between controller updates can offer greater design flexibility and increase the availability of computational resources. Having that in mind, a more suitable approach for scheduling controller updates would take into account sys-

tem states and perform updates only when necessary. Along these lines, the approaches of event- and self-triggering have recently been proposed to lengthen the intervals between the updates without sacrificing stability [1, 17–19, 21, 77, 104, 144, 162, 163, 172, 194, 202, 248, 275, 281–283]. The time at which the control should next be updated is derived in a way that ensures asymptotic stability of the closed-loop feedback control system with respect to errors caused by outdated samples. Since the decision to update hinges upon a state-dependent or sample-dependent criterion, the intervals between updates vary and may be longer than those under a periodic control implementation [21].

In an event-triggered control implementation, the system state is updated when it deviates from the previous sample by a sufficient amount [1, 21, 104, 172, 194, 202, 248, 281]. This requires continuous monitoring of the system state, which may be impractical. Therefore, in this paper, we consider a self-triggering approach in which the next update is defined based on the last sampled data [17, 18, 144, 162, 163, 275, 282]. In this case, the decision to update is computed from *predictions* of when the system state will deviate by a given threshold from the last update. However, for systems under the influence of disturbances or noise, it may be more difficult to make these predictions or to guarantee that the intended stability results are retained. Along these lines, the robustness of a self-triggered control strategy to disturbances was analyzed in [163] for linear systems. In [282], a self-triggered $\mathcal{H}_\infty$ control was developed for linear systems with a state-dependent disturbance, and this was extended in [283] for an exogenous

disturbance in $\mathcal{L}_2$ space.

In this work, we develop a self-triggered scheme for the control of stochastic dynamical systems described by stochastic differential equations (SDEs) [157] with additive noise terms. In the area of control, this type of dynamical model is commonly used for Kalman, or Extended Kalman filter designs [108]. For systems of this type, not only can it be difficult to predict the system state at a future time, but the noise may drastically alter the system dynamics. Because of that, we use the dynamics of certain statistics of the state distribution to develop an update rule that guarantees p-moment stability [116, 157] ($p > 0$) of the SDE solution. The length of times between controller updates is based on the previously-measured state and is shown to be strictly positive. Since the intensity of the noise in many stochastic systems is independent of the state, we split our analysis into two parts, one in which the noise diminishes as the system state approaches a stable equilibrium, and the other in which this is not the case.

In this paper, we simplify the update rule that was initially presented in [13] and better elucidate both the reasons for, and the consequences of the differences between our self-triggered approach and that for deterministic systems. We recently became aware of one other attempt at self-triggered stabilization of stochastic control systems. In [3], the authors develop a self-triggering rule for SDEs that guarantees stability in probability (note that this is weaker than p-moment stability). However, since their approach is based on a technique found in literature for event- and self-triggered deterministic systems, the assumptions on the control system are strict, and, as a result, no examples

have so far been provided with which we can compare the length of times between task updates. Therefore, another contribution of this paper is in providing examples that will be useful for future work comparisons. The first is a stochastic version of a linear control problem from deterministic literature [248], and the second is a stochastic wheeled cart control problem whose motivation is detailed in [11, 15, 16].

This paper is organized as follows: In Section 7.2, we introduce relevant notation and definitions, formulate the problem, and call attention to why the standard deterministic approach for event- and self-triggering may not apply, both theoretically and pragmatically, to stochastic systems. Section 7.3 presents some stability results for stochastic differential equations. Section 7.4 proposes a self-triggering scheme with strictly positive times between control updates, and Section 7.5 provides numerical examples. Finally, Section 7.6 summarizes the results of this paper and provides possible directions for future research.

7.2 Preliminaries and Problem Statement

7.2.1 Notation and Definitions

We will consider control systems defined by stochastic differential equations of the form

$$\mathrm{d}x(t) = f(x,u)\mathrm{d}t + g(x,u)\mathrm{d}\mathbf{w}, \qquad x \in \mathbb{R}^n, \tag{7.1}$$

where dw is a (multi-dimensional) increment of a standard Wiener process, $u(t) : [0,\infty) \to \mathbb{R}^m$ is a control input, and $f(\cdot,\cdot)$ and $g(\cdot,\cdot)$ are the drift and diffusion scaling factors of the dynamics. The differential operator $\mathcal{L}$ associated with a system in this form, when applied to a function $V(x,t)$ that is twice-differentiable in its first argument, is

$$\mathcal{L}V(x,t) = \frac{\partial V}{\partial t} + f^{\mathsf{T}}\frac{\partial V}{\partial x} + \frac{1}{2}\text{Trace}\left(g^{\mathsf{T}}\frac{\partial^2 V}{\partial x^2}g\right). \tag{7.2}$$

Usually these systems are formally defined alongside a complete probability space $(\Omega, \mathcal{F}, P)$ [183], where Ω is the set of possible outcomes, $\mathcal{F}$ is a filtration, and P is a probability measure function mapping $\mathcal{F} \to [0,1]$. We will also need the following definitions and notations. A function $\gamma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class $\mathcal{K}$ if it is continuous, strictly increasing, and $\gamma(0) = 0$. It is of class $\mathcal{K}_\infty$ if, in addition, it is unbounded. A function $\beta : \mathbb{R}_{\geq 0} \times [0,\infty) \to \mathbb{R}_{\geq 0}$ is of class $\mathcal{KL}$ if, for each fixed t, $\beta(x,t)$ is of class $\mathcal{K}$, and, for each fixed x, $\beta(x,t)$ is decreasing with $\beta(x,t) \to 0$ as $t \to \infty$. Let $|\cdot|$ denote the Euclidean norm $|x| = \sqrt{x^{\mathsf{T}}x}$. For two constants a and b, we use the notation $a \vee b = \max\{a,b\}$ and $a \wedge b = \min\{a,b\}$.

Definition 7.2.1 (cf. [152]) *A system is said to be* p-moment[1] input-to-state stable (ISS) *with respect to an error* $e(t)$ *if there exist a class* $\mathcal{KL}$ *function* β *and class* $\mathcal{K}$ *functions* γ *and* λ *such that for all* $t \geq 0$*,*

$$\mathbb{E}\left(|x(t)|^p\right) \leq \beta\left(\mathbb{E}\left(|x(0)|^p\right), t\right) + \gamma\left(\mathbb{E}\left(\lambda\left(\sup_{t\geq 0}|e(t)|\right)\right)\right). \tag{7.3}$$

[1] To be concise, note that $p-$th moment actually refers to the expected value of the Euclidean norm raised to the p-th power [157].

In many systems, the diffusion scaling $g(\cdot,\cdot)$ is independent of the state and control and does not vanish at the origin. For systems of this type, the second-order term in (7.2) may be a constant. In this work, we treat that constant as an additional disturbance, denoted d, that acts alongside the error due to sampling $e(t)$. The following definition will later be used to account for this.

Definition 7.2.2 (cf. [152]) *A system is said to be* practically p-moment stable *if there exist a class $\mathcal{KL}$ function β, a class $\mathcal{K}$ function γ, and a constant $d \geq 0$ such that for all $t \geq 0$,*

$$\mathbb{E}\left(|x(t)|^p\right) \leq \beta\left(\mathbb{E}\left(|x(0)|^p\right), t\right) + \gamma(d).$$

If $d = 0$, the system is said to be p-moment stable *(see Definition 2.1 in [116]).*

Definition 7.2.3 ([152]) *A function V is said to be a* stochastic input-to-state stable *Lyapunov function with respect to an error $e(t)$ if there exist class $\mathcal{K}_\infty$ functions $\underline{\alpha}$ and $\overline{\alpha}$ and class $\mathcal{K}$ functions α and λ such that*

$$\underline{\alpha}(|x|) \leq V(x) \leq \overline{\alpha}(|x|) \tag{7.4}$$

$$\mathcal{L}V \leq \lambda(|e|) - \alpha(|x|). \tag{7.5}$$

7.2.2 Problem Statement

We consider the state-feedback controlled system (7.1) with sample-and-hold state measurements, i.e.,

$$\mathrm{d}x(t) = f(x,u)\mathrm{d}t + g(x,u)\mathrm{d}\mathbf{w} \tag{7.6}$$

$$u(t) = k(x_i), \qquad t \in [t_i, t_{i+1}) \tag{7.7}$$

where t_i, $i = 0,1,\ldots$, is a sequence of update, or triggering, times, and $x_i = x(t_i)$, $i = 0,1,\ldots$, is the corresponding sequence of measurements of the system state that is used to update the feedback control $k(x_i)$. Let us define the error signal $e(t)$ as

$$e(t) = x_i - x(t), \qquad t \in [t_i, t_{i+1})\,. \tag{7.8}$$

Then (7.6) is

$$\mathrm{d}x(t) = f(x,k(x+e))\mathrm{d}t + g(x,k(x+e))\mathrm{d}\mathbf{w} \tag{7.9}$$

To simplify notation in the sequel, we will write with a slight abuse of notation $f(x,e)$ instead of $f(x,k(x+e))$ and $g(x,e)$ instead of $g(x,k(x+e))$. We make the following assumptions of monotone growth and local Lipschitz continuity, which are no stronger than the standard conditions for the existence and uniqueness of the process $x(t)$ (cf. [157, Theorem 2.3.5] with $e = 0$).

Assumption 1) There exists a positive constant K such that for all $x, e \in \mathbb{R}^n$ and $t \in [t_i, t_{i+1})$, $i = 0,1,\ldots$,

$$x^{\intercal} f(x,e) + \frac{1}{2}|g(x,e)|^2 \leq K(1 + |x|^2 + |e|^2) \tag{7.10}$$

Assumption 2) For every integer $m \geq 1$, there exist a positive constant L_m such that for all $x, e, x', e' \in \mathbb{R}^n$ with $|x| \vee |x'| \vee |e| \vee |e'| \leq m$ and $t \in [t_i, t_{i+1})$,

$$|f(x,e) - f(x',e')|^2 \vee |g(x,e) - g(x',e')|^2 \leq L_m(|x-x'|^2 + |e-e'|^2) \tag{7.11}$$

The goal of this paper is to develop an update rule for the stochastic system (7.8)-(7.9) based on the observable state x_i, $i = 0, 1, \ldots$ that will render the system practically stable while guaranteeing strictly positive inter-execution times $\tau_i = t_{i+1} - t_i$, that is, there is some minimum time between sampling time points. If $g(0,0) = 0$, i.e., if g admits an equilibrium point and the noise vanishes at the origin, then we will seek p-moment stability.

7.2.3 Motivating our Approach

For the purpose of motivating the sequel, we will first begin to formulate the problem using an approach found in literature for deterministic event-triggered and self-triggered control systems and describe how this fails when applied to stochastic control systems.

This usually begins with an assumption of input-to-state stability [235] of the closed-loop feedback control system with respect to errors caused by outdated samples. Along these lines, let us assume the existence of a stochastic input-to-state Lyapunov function

satisfying (7.4)-(7.5). If we were to further assume that the error were to satisfy [248]

$$\lambda(|e|) \leq \theta\alpha(|x|), \qquad 0 < \theta < 1, \tag{7.12}$$

then by (7.5), the state $x(t)$ is asymptotically stabilized since

$$\mathcal{L}V(x,t) \leq -(1-\theta)\alpha(|x|). \tag{7.13}$$

A condition of the form (7.12) can be used to implicitly define the sequence of times $\{t_i\}$. For example, a *triggering condition* of the form

$$\lambda(|e(t)|) = \theta\alpha(|x(t)|) \quad \Rightarrow \quad t_{i+1} \coloneqq t \tag{7.14}$$

would give rise to an event-triggered framework, since the update rule is based on the current state of the system $x(t)$ and the previous measurement x_i. If, additionally, predictions of the state $x(t)$ are available from the knowledge of x_i, one could extend (7.12) into a self-triggered scheme.

In order to show that task periods τ_i are bounded strictly away from zero, the standard technique in literature is to then determine the duration for which the error $e(t)$ satisfies a condition like (7.12) [248, 282], either using the value of the state (for an event-triggered scheme) or using predictions (a self-triggered scheme). However, in the stochastic case considered in this work, the error may exceed this bound instantaneously, that is, for any $M < \infty$ and any time $t > 0$, the Euclidean norm of a solution $e(t)$ to a stochastic differential equation will exceed the level M with non-zero probability, or $\Pr(|e(t)| \geq M) > 0$ [183, Exercise 8.13]. In other words, although certain trajectories of $x(t)$ and $e(t)$ for fixed $\omega \in \Omega' \subset \Omega$ may satisfy the desired triggering

criteria for a sufficiently large time, the same can not be said about all trajectories $x(t)$ and $e(t)$ defined by (7.8)-(7.9). Moreover, the second-order differential terms required in stochastic evolution equations can cause certain quantities found in deterministic literature (e.g., the quantity $|e(t)|/|x(t)|$ [248]) to experience unbounded growth near the origin. These facts make it difficult to develop a sampling rule using the trajectories $e(t)$ and $x(t)$ or predictions of these processes. Because of this, we instead consider the p-th moments of these processes, $\mathbb{E}(|e|^p)$ and $\mathbb{E}(|x|^p)$, and develop a triggering condition based on these statistics to guarantee the stability of $x(t)$ in the p-th moment [157].

As a consequence of this basis for the update rule, we must rule out an event-triggered implementation. From a practical standpoint, the controller can only measure an individual sample path of the process $x(t)$, and not the statistics $\mathbb{E}(|e|^p)$ or $\mathbb{E}(|x|^p)$. However, the latter quantities can be predicted on the interval $[t_i, t_{i+1})$ based on the last-sampled state x_i, and are therefore suitable for a self-triggered approach.

7.3 A Lyapunov Characterization for p-moment Input-to-state Stability

Since this work deals with stability in the p-th moment, we must first revise the preliminary notion of stability used to motivate this work (7.4)-(7.5) and provide a Lyapunov function-based stability criterion for p-moment ISS. The characterization of stability in the following theorem is similar to that provided in (7.4)-(7.5), but with the addition of

expectation operators so that we may examine triggering rules based on the statistics $\mathbb{E}(|x|^p)$ and $\mathbb{E}(|e|^p)$ of the processes considered in this work.

Theorem 7.3.1 *Suppose there exist a convex class $\mathcal{K}_\infty$ function $\underline{\alpha}$, a class $\mathcal{K}_\infty$ function $\overline{\alpha}$, a non-negative function α, and non-negative function $V(x,t)$ that is twice differentiable in its first argument such that*

$$\underline{\alpha}(|x|^p) \leq V(x,t) \leq \overline{\alpha}(|x|^p), \tag{7.15}$$

$$\mathbb{E}\mathcal{L}V(x,t) \leq \mathbb{E}(\lambda(|e|)) - \mathbb{E}(\alpha(x(t))) \tag{7.16}$$

for all $t \geq 0$, where $\lim_{|x|\to\infty} \alpha(x)/\overline{\alpha}(|x|^p) > 0$. Then the system (7.6)-(7.8) is p-moment ISS.

Theorem 7.3.1 is a specific case of [116, Theorem 3.1]. The latter applies to more general systems with delays and Markovian switching. That the Lyapunov characterization we use arises from the study of stochastic differential *delay* equations should not come as a surprise, since a sample x_i really amounts to a delayed state variable. However, in order to more closely tie these ideas with previous literature on event- and self-triggered control, we have removed the notion of a delay. Nevertheless, the proof of Theorem 7.3.1 is very similar to that of [116, Theorem 3.1], but for completeness, we include the relevant elements of that work here. We first provide an accompanying lemma.

Lemma 7.3.1 *For any $t \geq 0$, there is a constant $a_w > 0$ such that $\mathbb{E}(\alpha(x)) \geq a_w$ whenever $\mathbb{E}(V(x,t)) \geq a_v > 0$.*

Proof: We show the existence of a class $\mathcal{K}_\infty$ function $\mu_w(\cdot)$ such that $\mathbb{E}(\alpha(x)) \geq \mu_w(a_v) = a_w$ whenever $\mathbb{E}(\overline{\alpha}(|x|^p)) \geq a_v > 0$. Fix t and define a function $b : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$

$$b(y) = \inf_{|x|^p \geq \overline{\alpha}^{-1}(y/2)} \frac{\alpha(x)}{\overline{\alpha}(|x|^p)}, \qquad y \geq 0,$$

that is non-decreasing by construction. By the properties of $\alpha(x)$ and $\overline{\alpha}(|x|^p)$, we have that $b(y) > 0$ whenever $y > 0$. Then if, for any $a_v > 0$, $\mathbb{E}(\overline{\alpha}(|x|^p)) \geq \mathbb{E}(V(x,t))) \geq a_v$, we have that

$$\begin{aligned}
\mathbb{E}(\alpha(x)) &= \int_\Omega \alpha(x) dP \\
&\geq \int_{|x|^p \geq \overline{\alpha}^{-1}(a_v/2)} \alpha(x) dP \\
&\geq b(a_v) \int_{\overline{\alpha}(|x|^p) \geq a_v/2} \overline{\alpha}(|x|^p) dP \geq \frac{a_v b(a_v)}{2},
\end{aligned}$$

where the lemma follows with $\mu_w(a_v) = a_v b(a_v)/2$. □

We now continue with the proof of the theorem adapted from [116].

Proof: Let $e_\lambda = \sup_{t \geq 0} \mathbb{E}(\lambda(|e|)) \leq \mathbb{E}\left(\lambda(\sup_{t \geq 0} |e(t)|)\right)$ and $\underline{\alpha}_0 = \underline{\alpha}(\mathbb{E}(|x(0)|^p))$. Without loss of generality, choose $\mu_w(\cdot)$ such that $0 < \mu_w^{-1}(2e_\lambda) < \underline{\alpha}_0$. We will show that

$$\mathbb{E}(V(x,t)) \leq \underline{\alpha}_0 \wedge \mu_w^{-1}(2e_\lambda), \qquad t \geq 0. \tag{7.17}$$

We first show that $\mathbb{E}(V(x,t)) \leq \underline{\alpha}_0$ for all $t \geq 0$. Suppose that

$$t_a = \inf\{t > 0 : \mathbb{E}(V(x,t)) > \underline{\alpha}_0\},$$

and assume for the purpose of contradiction that $t_a < \infty$. Since $\mathbb{E}(V(x,t))$ is continuous,

there must be constants t_b, t_c, and $\varepsilon > 0$ satisfying $0 \leq t_b \leq t_a < t_c$ and

$$\begin{cases} \mathbb{E}(V(x,t)) = \underline{\alpha}_0 & t = t_b \\ \\ \underline{\alpha}_0 < \mathbb{E}(V(x,t)) < \varepsilon + \underline{\alpha}_0 & t_b < t \leq t_c \end{cases} \tag{7.18}$$

However, the expected value of the Lyapunov function satisfies

$$\begin{aligned} \mathbb{E}(V(x(t),t)) &= \mathbb{E}(V(x(t_b),t_b)) + \mathbb{E}\int_{t_b}^{t} \mathcal{L}V(x(s),s)ds \\ &= \mathbb{E}(V(x(t_b),t_b)) + \int_{t_b}^{t} \mathbb{E}\mathcal{L}V(x(s),s)ds \\ &\leq \underline{\alpha}_0 + \int_{t_b}^{t} (e_\lambda - 2e_\lambda)\, ds \leq \underline{\alpha}_0 - e_\lambda (t - t_b) < \underline{\alpha}_0 \end{aligned}$$

for $t_b < t \leq t_c$, which contradicts (7.18), so $\mathbb{E}(V(x,t)) \leq \underline{\alpha}_0$ for all $t \geq 0$.

Next, we show that for a time t_μ satisfying $t_\mu / e_\lambda > \underline{\alpha}_0$, we have $\mathbb{E}(V(x,t)) \leq \mu_w^{-1}(2e_\lambda)$ for all $t \geq t_\mu$. Let $\tau = \inf\left\{t \geq 0 : \mathbb{E}(V(x,t)) \leq \mu_w^{-1}(2e_\lambda)\right\}$ and suppose for the purpose of contradiction that $\tau > t_\mu$. Then $\mathbb{E}(V(x,t)) \geq \mu_w^{-1}(2e_\lambda)$ for $t \leq t_\mu$, and by Lemma 7.3.1, we have $\mathbb{E}(\alpha(x)) \geq 2e_\lambda$, so that for all $0 \leq t \leq t_\mu$,

$$\begin{aligned} \mathbb{E}(V(x(t),t)) &= \mathbb{E}(V(x(0),0)) + \int_{0}^{t} \mathbb{E}\mathcal{L}V(x(s),s)ds \\ &\leq \underline{\alpha}_0 + \int_{0}^{t\mu} (e_\lambda - 2e_\lambda) ds \leq \underline{\alpha}_0 - e_\lambda t_\mu < 0, \end{aligned} \tag{7.19}$$

which violates the fact that $\mathbb{E}(V(x,t)) \geq 0$ for all $t \geq 0$, and so $\tau \leq t_\mu$. Next, let $t_{\mu a} = \inf\left\{t > \tau : \mathbb{E}(V(x,t)) > \mu_w^{-1}(2e_\lambda)\right\}$ and suppose that $t_{\mu a} < \infty$. Then there must exist

constants $t_{\mu b}$ and $t_{\mu c}$ such that $t_\mu \leq t_{\mu b} \leq t_{\mu a} < t_{\mu c}$ and

$$\begin{cases} \mathbb{E}(V(x,t)) = \mu_w^{-1}(2e_\lambda) & t = t_{\mu b} \\ \mu_w^{-1}(2e_\lambda) < \mathbb{E}(V(x,t)) < \varepsilon + \mu_w^{-1}(2e_\lambda) & t_{\mu b} < t \leq t_{\mu c} \end{cases}$$

Consequently, we can derive in a similar manner that $\mathbb{E}(V(x(t),t)) \leq \mu_w^{-1}(2e_\lambda) - e_\lambda(t - t_{\mu b})$, which again contradicts the fact that $\mathbb{E}(V(x,t)) \geq 0$ for all $t \geq 0$, and so $\mathbb{E}(V(x,t)) \leq \mu_w^{-1}(2e_\lambda)$ for all $t \geq t_\mu$.

We return to the statement of the theorem to show that, using Jensen's inequality and (7.17), for $t \geq t_\mu$,

$$\begin{aligned} \underline{\alpha}\left(\mathbb{E}(|x|^p)\right) \leq \mathbb{E}\left(\underline{\alpha}(|x|^p)\right) &\leq \mathbb{E}(V(x,t)) \\ &\leq \mu_w^{-1}(2e_\lambda) \leq \mu_w^{-1}\left(2\mathbb{E}\left(\lambda\left(\sup_{t\geq 0}|e(t)|\right)\right)\right), \end{aligned} \tag{7.20}$$

which implies that

$$\begin{aligned} \mathbb{E}\left(|x|^p\right) &\leq \underline{\alpha}^{-1}\left(\mu_w^{-1}\left(2\mathbb{E}\left(\lambda\left(\sup_{t\geq 0}|e(t)|\right)\right)\right)\right) \\ &\equiv \gamma\left(\mathbb{E}\left(\lambda\left(\sup_{t\geq 0}|e(t)|\right)\right)\right), \qquad t \geq t_\mu, \end{aligned} \tag{7.21}$$

where $\gamma(|x|) = \underline{\alpha}^{-1}\left(\mu_w^{-1}\left(2|x|\right)\right)$ is class $\mathcal{K}$ by construction.

Let $\widetilde{\beta}$ be a class $\mathcal{KL}$ function satisfying $\widetilde{\beta}(\underline{\alpha}(|x_0|^p,t) = \widetilde{\beta}(\underline{\alpha}_0,t) \geq 2\underline{\alpha}_0 - \underline{\alpha}_0(t/t_\mu)$ for all $0 \leq t \leq t_\mu$. Then it follows from (7.17) that $\mathbb{E}(V(x,t)) \leq \widetilde{\beta}(\underline{\alpha}_0,t)$ for all $0 \leq t \leq t_\mu$, and consequently

$$\mathbb{E}\left(|x|^p\right) \leq \underline{\alpha}^{-1}\left(\widetilde{\beta}(\underline{\alpha}_0,t)\right) = \beta\left(\mathbb{E}(|x(0)|^p),t\right) \tag{7.22}$$

for $0 \leq t \leq t_\mu$, where $\beta(x,t) = \underline{\alpha}^{-1}\left(\widetilde{\beta}(\underline{\alpha}(x),t)\right)$ is a class $\mathcal{KL}$ function. The desired result (7.3) follows from (7.21)-(7.22). □

7.4 Triggering Condition

Suppose that there is a Lyapunov function for the system (7.6)-(7.8) satisfying (7.15) and

$$\mathbb{E}\mathcal{L}V(x,t) \leq \mathbb{E}(\lambda(|e|)) - \mathbb{E}(\alpha(x(t))) + d \tag{7.23}$$

Here, $d > 0$ allows for the possibility of a constant disturbance that does not diminish at the origin. This may occur, for example, if the second-order term in (7.2) is a constant. In our approach, we use this secondary disturbance to our advantage and include it in the update rule. Next, suppose that the error were to satisfy (cf. (7.12))

$$\mathbb{E}(\lambda(|e|)) \leq \theta\mathbb{E}(\alpha(x)) + \theta_d d, \tag{7.24}$$

for a constant $0 < \theta < 1$ and parameter $\theta_d > 0$. Then from (7.16),

$$\mathbb{E}\mathcal{L}V \leq -(1-\theta)\mathbb{E}(\alpha(x)) + (1+\theta_d)d. \tag{7.25}$$

In this case, (7.25) does not include the error $e(t)$ due to the sampling rule, but does include the constant disturbance $(1+\theta_d)d$. By rewriting (7.16) with $\lambda(|e|)$ replaced by $(1+\theta_d)d$ and using Theorem 7.3.1, one can show that the system will be practically p-moment stable with $\gamma(d)$ in (7.3) replaced by $\gamma((1+\theta_d)d)$. If $(1+\theta_d)d = 0$, i.e., if $d = 0$, then the system will be p-moment stable.

7.4.1 Predictions of the moments of the processes

Before developing our controller update rule, we require some relations that can be derived from the SDE assumptions (7.10)-(7.11). These lemmas relate the predictions of the statistics $\mathbb{E}(|e|^2)$ and $\mathbb{E}(|x|^2)$, which are not observable, to the norm of the last-observed system state x_i.

Lemma 7.4.1 *Assume the monotone growth assumption (7.10). If the system state has been updated as $x_i = x(t_i)$, then for any $t \in [t_i, t_{i+1})$, the means of the norms $|e(t)|^2$ and $|x(t)|^2$ are upper and lower bounded, respectively, based on the following inequalities*

$$\mathbb{E}(|e(t)|^2) \leq A(|x_i|, t-t_i) \tag{7.26}$$

$$\mathbb{E}(|x(t)|^2) \geq B(|x_i|, t-t_i) \tag{7.27}$$

where

$$A(|x_i|, t-t_i) = \frac{2|x_i|^2+1}{3}\left(e^{12K(t-t_i)}-1\right) \tag{7.28}$$

$$B(|x_i|, t-t_i) = \frac{5|x_i|^2+1}{3}e^{-6K(t-t_i)} - \frac{2|x_i|^2+1}{3} \tag{7.29}$$

Proof: Using Itô's Lemma and (7.8), we get

$$\begin{aligned}
\mathrm{d}\mathbb{E}(|e|^2) &\leq \mathrm{d}\mathbb{E}\left(2|x_i|^2+2|x|^2\right) \\
&= 2\mathbb{E}\left(2x^\intercal f(x,e)\mathrm{d}t + 2x^\intercal g(x,e)\mathrm{d}w + |g(x,e)|^2\mathrm{d}t\right) \\
&= 4\mathbb{E}\left(x^\intercal f(x,e) + \frac{1}{2}|g(x,e)|^2\right)\mathrm{d}t
\end{aligned}$$

From the monotone growth condition (7.10)

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(|e|^2) &\leq 4K\left(1+\mathbb{E}(|x|^2)+\mathbb{E}(|e|^2)\right)\\ &\leq 4K\left(1+\mathbb{E}(|x_i-e|^2)+\mathbb{E}(|e|^2)\right)\\ &\leq 4K\left(1+3\mathbb{E}(|e|^2)+2\mathbb{E}(|x_i|^2)\right)\\ &\leq 12K\mathbb{E}(|e|^2)+8K|x_i|^2+4K\end{aligned}$$

Applying the comparison principle in [130], along with the fact that $e(t_i)=0$, we obtain (7.26). For the second inequality (7.27), we can obtain in a similar manner

$$\begin{aligned}\left|\mathrm{d}\mathbb{E}(|x|^2)\right| &\leq \left|\mathbb{E}\left(2x^{\mathsf{T}}f(x,e)\mathrm{d}t+2x^{\mathsf{T}}g(x,e)\mathrm{d}w+|g(x,e)|^2\mathrm{d}t\right)\right|\\ &= 2\left|\mathbb{E}\left(x^{\mathsf{T}}f(x,e)+\frac{1}{2}|g(x,e)|^2\right)\mathrm{d}t\right|\\ &\leq 6K\mathbb{E}(|x|^2)\mathrm{d}t+4K|x_i|^2\mathrm{d}t+2K\mathrm{d}t,\end{aligned}$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(-\mathbb{E}(|x|^2)\right)\leq -6K\left(-\mathbb{E}(|x|^2)\right)+4K|x_i|^2+2K. \tag{7.30}$$

Using the comparison principle with $x(t_i)=x_i$ yields (7.27). □

Lemma 7.4.2 *Assume the local Lipschitz continuity assumption (7.11) where $f(\cdot,\cdot)$ and $g(\cdot,\cdot)$ admit an equilibrium point, i.e., $f(0,0)=0$ and $g(0,0)=0$. If the system state has been updated as $x_i=x(t_i)$, then for any $t\in[t_i,t_i+\tau_i)$, the quantities $\mathbb{E}(|e(t)|^2)$ and $\mathbb{E}(|x(t)|^2)$ are upper and lower bounded, respectively, based on the following in-*

equalities

$$\mathbb{E}(|e(t)|^2) \leq C(|x_i|, t-t_i) \tag{7.31}$$

$$\mathbb{E}(|x(t)|^2) \geq D(|x_i|, t-t_i) \tag{7.32}$$

where

$$C(|x_i|, t-t_i) = \frac{2|x_i|^2(1+\sqrt{L_m})}{4+3\sqrt{L_m}} \left(e^{(4\sqrt{L_m}+3L_m)(t-t_i)} - 1\right) \tag{7.33}$$

$$D(|x_i|, t-t_i) = \frac{|x_i|^2}{4+3\sqrt{L_m}} \left((6+5\sqrt{L_m})e^{-(4\sqrt{L_m}+3L_m)(t-t_i)} - 2(1+\sqrt{L_m})\right) \tag{7.34}$$

Proof: Recall that from (7.6) and (7.8), the error kinematics satisfy for $t \in [t_i, t_{i+1})$:

$$e(t) = -\int_{t_i}^{t} f(x,u)\mathrm{d}s - \int_{t_i}^{t} g(x,u)\mathrm{d}\mathbf{w}_s. \tag{7.35}$$

Using Itô's Lemma and the inequality $2ab \leq a^2 + b^2$,

$$\begin{aligned}
\mathrm{d}\mathbb{E}(|e|^2) &= \mathbb{E}\left(-2e^{\mathsf{T}} f(x,e)\mathrm{d}t - 2e^{\mathsf{T}} g(x,e)\mathrm{d}w + |g(x,e)|^2\mathrm{d}t\right) \\
&= -2\mathbb{E}\left(e^{\mathsf{T}} f(x,e)\right)\mathrm{d}t + \mathbb{E}(|g(x,e)|^2)\mathrm{d}t \\
&\leq 2\mathbb{E}\left(\left[L_m^{\frac{1}{4}}|e|\right]\left[\frac{|f(x,e)|}{L_m^{\frac{1}{4}}}\right]\right)\mathrm{d}t + \mathbb{E}(|g(x,e)|^2)\mathrm{d}t \\
&\leq \sqrt{L_m}\mathbb{E}\left(|e|^2\right)\mathrm{d}t + \frac{1}{\sqrt{L_m}}\mathbb{E}\left(|f(x,e)|^2\right)\mathrm{d}t + \mathbb{E}(|g(x,e)|^2)\mathrm{d}t
\end{aligned} \tag{7.36}$$

Inserting the Lipschitz condition (7.11) with $x' = e' = 0$, we obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(|e|^2) &\leq \sqrt{L_m}\mathbb{E}\left(|e|^2\right) + (\sqrt{L_m}+L_m)\mathbb{E}\left(|x_i - e|^2 + |e|^2\right) \\
&\leq (4\sqrt{L_m}+3L)\mathbb{E}(|e|^2) + 2(\sqrt{L_m}+L_m)|x_i|^2
\end{aligned}$$

From the comparison principle and the fact that $e(t_i) = 0$, we obtain (7.31). The second inequality (7.32) follows similarly. □

7.4.2 A Self-triggering Sampling Rule

With these inequalities on $\mathbb{E}(|e(t)|^2)$ and $\mathbb{E}(|x(t)|^2)$, we are ready to state our main results. The following theorems provide relations based on (7.24) that can be used to calculate a strictly positive inter-execution time $\tau_i = t_{i+1} - t_i$ as a function of the last-observed state x_i. The first theorem requires Lemma 7.4.1, which does not assume that $g(0,0) = 0$, and guarantees practical p-th moment stability ($d > 0$). The second theorem uses Lemma 7.4.2, which assumes a state-dependent diffusion coefficient such that $g(0,0) = 0$, and allows for $d \geq 0$, i.e., p-moment stability can be achieved with $d = 0$.

Theorem 7.4.1 *Assume that in addition to the conditions of Theorem 7.3.1 and Lemma 7.4.1, there exist a convex class $\mathcal{K}$ function $\alpha_v(\cdot)$ and a concave class $\mathcal{K}$ function $\lambda_c(\cdot)$ which satisfy*

$$\alpha_v(2|x|^2) \leq 2\alpha(x) \tag{7.37}$$

$$\lambda_c(|e|) \geq \lambda(\sqrt{|e|}). \tag{7.38}$$

Suppose that the system (7.6)-(7.8) has been updated at $t = t_i$ with state x_i, and that the time until the next update $\tau_i = t_{i+1} - t_i$ is such that

$$\lambda_c\left(A(|x_i|, \tau_i)\right) \leq \theta \alpha_v\left(B(|x_i|, \tau_i) + d_\alpha\right). \tag{7.39}$$

Then (7.24) will hold with

$$\theta_d d = \theta \alpha_v(2d_\alpha)/2, \tag{7.40}$$

ensuring practical p-moment stability. Moreover, if $d_\alpha > 0$ and $|x_i| \leq \bar{x} < \infty$, the execution times do not reach an accumulation point, i.e., $\tau_i > 0$, $i = 0, 1, \ldots$.

Proof: Substitution of the inequalities (7.26) and (7.27) into (7.39) gives

$$\lambda_c\left(\mathbb{E}(|e|^2)\right) \leq \theta\alpha_v\left(\mathbb{E}\left(|x|^2\right) + d_\alpha\right). \tag{7.41}$$

Since $\alpha_v(\cdot)$ is convex, we have from (7.37) and Jensen's inequality that the right hand side of (7.41) is

$$\begin{aligned}
&\theta\alpha_v\left(\mathbb{E}\left(|x|^2\right) + d_\alpha\right) \\
&\quad = \theta\alpha_v\left(2\left(\frac{1}{2}\mathbb{E}(|x|^2) + \frac{1}{2}d_\alpha\right)\right) \\
&\quad \leq \theta\mathbb{E}\left(\alpha_v(2|x|^2)/2\right) + \theta\alpha_v(2d_\alpha)/2 \\
&\quad \leq \theta\mathbb{E}\left(\alpha(|x|)\right) + \theta_d d
\end{aligned} \tag{7.42}$$

which is the right hand side of (7.24) with $\theta_d d = \theta\alpha_v(2d_\alpha)/2$. Similarly, through the concavity of $\lambda_c(\cdot)$ and (7.38), the left hand side of (7.41) can also be made to match that of (7.24) using the assumption $\lambda_c\left(\mathbb{E}(|e|^2)\right) \geq \mathbb{E}\left(\lambda(|e|)\right)$:

$$\mathbb{E}\left(\lambda(|e|)\right) \leq \mathbb{E}\left(\lambda_c(|e|^2)\right) \leq \lambda_c\left(\mathbb{E}(|e|^2)\right).$$

In summary, we have shown that (7.39) implies (7.24), i.e.,

$$\mathbb{E}\left(\lambda(|e|)\right) \leq \lambda_c\left(A(|x_i|, \tau_i)\right) \leq \theta\alpha_v\left(B(|x_i|, \tau_i) + d_\alpha\right) \leq \theta\mathbb{E}\left(\alpha(|x|)\right) + \theta_d d.$$

Next, to show the existence of a lower bound for the inter-execution times implicitly

defined by (7.39), let us define a function $\kappa(|x_i|)$ that is chosen to satisfy

$$0 < \kappa(|x_i|) < |x_i|^2 + d_\alpha. \qquad \text{for all } |x_i| \geq 0 \tag{7.43}$$

Clearly such a $\kappa(\cdot)$ exists, and, moreover, $\kappa(|x_i|) > A(|x_i|,0)$ and $\kappa(|x_i|) < B(|x_i|,0) + d_\alpha$. We need to verify that there exists a τ_i such that

$$\begin{aligned}\lambda_c\left(A(|x_i|,0)\right) &< \lambda_c\left(\kappa(|x_i|)\right) \leq \lambda_c\left(A(|x_i|,\tau_i)\right) \\ &\leq \theta\alpha_v\left(B(|x_i|,\tau_i)+d_\alpha\right) \leq \alpha_v\left(\kappa(|x_i|)\right) < \alpha_v\left(B(|x_i|,0)+d_\alpha\right).\end{aligned}$$

By solving for τ_i in both $\kappa(|x_i|) \leq A(|x_i|,\tau_i)$ and $\kappa(|x_i|) \geq B(|x_i|,\tau_i) + d_\alpha$, we obtain

$$\begin{aligned}\tau_i \geq \frac{1}{12K}\ln\left(1+\kappa(|x_i|)\frac{3}{2|x_i|^2+1}\right) \\ \wedge\ -\frac{1}{6K}\ln\left(1-\frac{3(|x_i|^2+d_\alpha-\kappa(|x_i|))}{5|x_i|^2+1}\right) > 0.\end{aligned} \tag{7.44}$$

where the strict positivity of the right hand side is due to the relation (7.43). □

Remark 7.4.1 *The inter-execution times τ_i, $i = 0,1,\ldots$, may be calculated numerically from (7.39) based on the norm of the last-observed state $|x_i|$.*

Remark 7.4.2 *We have not assumed that $g(0,0) = 0$ for the proof of this theorem. Therefore, this result is useful for systems where noise does not vanish at the origin. In this case, it is likely that $d > 0$ (and $d_\alpha > 0$).*

Note that for some systems the duration of the inter-execution times τ_i defined by (7.39) may be lengthened with increasing d_α. However, with longer inter-execution times due to increased d_α comes a larger value of $\gamma((1+\theta_d)d)$ in (7.3).

If the constant disturbance is $d = 0$ (which is often the case when $g(0,0) = 0$), one could apply Theorem 7.4.1 with a $d_\alpha > 0$, although this would only guarantee practical p-th moment stability, and it would introduce a $d > 0$ to (7.24). To avoid this, we next show how p-moment stability (i.e., $d = 0$) can be achieved for systems where $f(0,0) = g(0,0) = 0$ in the following theorem.

Theorem 7.4.2 *Assume that in addition to the conditions of Theorem 7.3.1 and Lemma 7.4.2, there exist a convex class $\mathcal{K}$ function $\alpha_v(\cdot)$ and a concave class $\mathcal{K}$ function $\lambda_c(\cdot)$ as in (7.37)-(7.38). Suppose that the system (7.6)-(7.8) has been updated at $t = t_i$ with state x_i, $|x_i| > 0$, and that the time until the next update $\tau_i = t_{i+1} - t_i$ is such that*

$$\lambda_c\left(C(|x_i|, \tau_i)\right) \leq \theta \alpha_v\left(D(|x_i|, \tau_i) + d_\alpha\right) \tag{7.45}$$

Then (7.24) will hold with $\theta_d d = \theta \alpha_v(2d_\alpha)/2$ (and $d = 0$ if $d_\alpha = 0$), ensuring practical p-moment stability if $d_\alpha > 0$ or p-moment stability if $d_\alpha = 0$. Moreover, for any non-negative d_α and $|x_i| \leq \bar{x} < \infty$, the execution times do not reach an accumulation point, i.e., $\tau_i > 0$, $i = 0, 1, \ldots$.

Proof: The desired relation (7.24) follows in the same way as in the proof of Theorem 7.4.1 with the use of $C(|x_i|, \tau_i)$ and $D(|x_i|, \tau_i)$ in place of $A(|x_i|, \tau_i)$ and $B(|x_i|, \tau_i)$. For the lower bound on τ_i, choose a function $\kappa(|x_i|)$ to satisfy

$$0 < \kappa_\lambda\left(|x_i|\right) < |x_i|^2 + d_\alpha, \qquad \text{for all } |x_i| > 0 \tag{7.46}$$

$$\lim_{|x_i| \to 0^+} \frac{\kappa(|x_i|)}{|x_i|^2} < 1 \tag{7.47}$$

Inverting $\kappa(|x_i|) \leq C(|x_i|, \tau_i)$ and $\kappa(|x_i|) \geq D(|x_i|, \tau_i)$ for τ_i, we obtain

$$\tau_i \geq \frac{1}{4+3\sqrt{L_m}} \ln\left(1 + \frac{4+3\sqrt{L_m}}{2+2\sqrt{L_m}} \frac{\kappa(|x_i|)}{|x_i|^2}\right) \\ \wedge\ -\frac{1}{4+3\sqrt{L_m}} \ln\left(1 - \frac{4+3\sqrt{L_m}}{6+5\sqrt{L_m}} \frac{(|x_i|^2 + d_\alpha - \kappa(|x_i|))}{|x_i|^2}\right) \tag{7.48}$$

and using (7.46)-(7.47), $\tau_i > 0$. □

Remark 7.4.3 *Note that (7.45) must only hold for $|x_i| > 0$ while (7.39) assumes $|x_i| \geq 0$. This is the case for two reasons. First, Theorem 7.4.2 applies to systems admitting the trivial solution $x(t) = 0$, and so the case $|x_i| = 0$ can be excluded. Moreover, choose in the proof of Theorem 7.4.2 $\kappa(|x_i|) = a|x_i|^2$ for $0 < a < 1$, for example, and let $d_\alpha = 0$. Then (7.48) becomes*

$$\tau_i \geq \frac{1}{4+3\sqrt{L_m}} \ln\left(1 + a\frac{4+3\sqrt{L_m}}{2+2\sqrt{L_m}}\right) \\ \wedge\ -\frac{1}{4+3\sqrt{L_m}} \ln\left(1 - (1-a)\frac{4+3\sqrt{L_m}}{6+5\sqrt{L_m}}\right) > 0.$$

7.5 Numerical Examples

In this section we provide examples and show how the triggering rule should change depending on the form of the intensity of noise at $x = 0$.

7.5.1 Stochastic Linear System

The first example is drawn from [248]. In the example, we have added Wiener process increments with both a constant scaling factor $\sigma/2$ and a state-dependent coefficient

$\sigma_x|x|/2$. The system is

$$\mathrm{d}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\mathrm{d}t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u\mathrm{d}t + \frac{1}{2}(\sigma + \sigma_x|x|)\begin{bmatrix} \mathrm{d}w_1 \\ \mathrm{d}w_2 \end{bmatrix}$$

with $u = x_1 - 4x_2$. Using $V(x) = x^\intercal Px$ as a Lyapunov function, we can obtain $\mathcal{L}V \leq -x^\intercal Qx + \sigma_x^2|x|^2 + \sigma^2$ with

$$P = \begin{bmatrix} 1 & \frac{1}{4} \\ \frac{1}{4} & 1 \end{bmatrix}, \qquad Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{2} \end{bmatrix}$$

Under a sampled-data implementation, this becomes

$$\mathcal{L}V \leq -(a - \sigma_x^2)|x|^2 + b|e||x| + \sigma^2$$

where $a = \lambda_m(Q) > 0.44$ is the smallest eigenvalue of Q, and $b = |K^\intercal B^\intercal P + PBK| = 8$. As our triggering results involve expectation operators, we take the expectation of both sides and apply Hölder's inequality

$$\begin{aligned}
\mathbb{E}\mathcal{L}V &\leq -(a - \sigma_x^2)\mathbb{E}|x|^2 + b\mathbb{E}(|e||x|) + \sigma^2 \\
&\leq -(a - \sigma_x^2)\mathbb{E}|x|^2 + b\sqrt{\mathbb{E}(|e|^2)}\sqrt{\mathbb{E}(|x|^2)} + \sigma^2 \\
&\leq -(a - \sigma_x^2)\mathbb{E}|x|^2 + b\sqrt{\mathbb{E}(|e|^2)}\sqrt{\mathbb{E}(|x|^2) + \frac{\theta_d}{\theta}\sigma^2} + \sigma^2
\end{aligned}$$

where the constant term involving $\theta_d > 0$ has been added in order to facilitate the following triggering rule. If we were to assume that

$$\mathbb{E}(|e|^2) \leq \theta\mathbb{E}(|x|^2) + \theta_d\sigma^2, \tag{7.49}$$

for some constant $0 < \theta < (a - \sigma_x^2)^2/b^2$, then

$$\mathbb{E}\mathcal{L}V \leq -(a - \sigma_x^2 - b\sqrt{\theta})\mathbb{E}|x|^2 + \sigma^2(1 + \frac{\theta_d}{\sqrt{\theta}}b)$$

and based on Theorem 7.3.1, we will obtain practical stability in the $p = 2$ moment, i.e., in the mean square sense. In light of the triggering rule (7.49), we set[2] $\alpha_v(|x|) = |x|$, and $\lambda_c(|e|) = |e|$, meaning that the update times τ_i given $|x_i|$ can be solved numerically using Theorem 7.4.1 from

$$A(|x_i|, \tau_i) \leq \theta B(|x_i|, \tau_i) + d_\alpha \tag{7.50}$$

with $d_\alpha = \theta_d \sigma^2/\theta$. Note that with these choices of $\lambda_c(|e|)$ and $\alpha_v(|x|)$, it is possible to write down an analytic triggering rule through an appropriate choice of $\kappa(|x_i|)$ using (7.43) and (7.44), but this may decrease the duration of the time between updates.

For our first simulations of this system, we choose $\sigma = \sigma_x = 0.1$. The monotone growth and Lipschitz coefficients in (7.10) and (7.11) are

$$K = \left(\frac{|BK|^2}{2} + \frac{\sigma_x^2}{4} - 1 \right) \vee \frac{|BK|}{2} \vee \frac{\sigma^2}{4} = 7.5025$$

$$L_m = (2\sigma_x^2 + 2|A + BK|^2) \vee 2|BK|^2 = 34,$$

respectively, and $(a - \sigma_x^2)/b = 0.0537$, so we choose $\theta = 0.0028$ and $\theta_d = 0.28$ (so that $d_\alpha = 1$).

Fig. 7.1(a) shows $\mathbb{E}(|x|^2)$ based on 1000 simulations from an initial condition on a random vector of magnitude $|x(0)| = 5$. During simulation, after the state has been measured as x_i, Theorem 7.4.1 provides a *deterministic* time $t = t_i + \tau_i$ at which to update the state. However, with each simulation, we obtain different samples x_i, $i = 1, \ldots$, and, consequently, the values of τ_i are random when not conditioned on x_i. Fig. 7.1(b) shows

[2]Recall that only concavity and convexity, and not strict concavity and convexity, are assumed for $\lambda_c(\cdot)$ and $\alpha_v(\cdot)$, respectively.

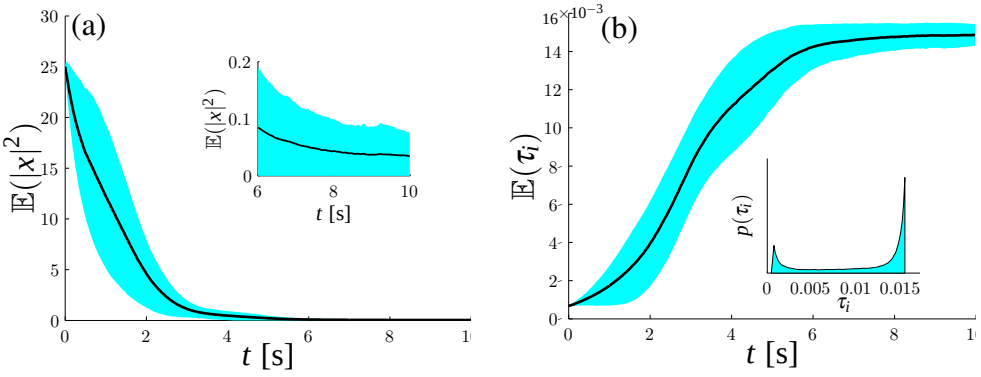


Figure 7.1: Linear system from [248] with stochasticity added. (a) Evolution of $\mathbb{E}(|x|^2)$ over a 10 s simulation, averaged over 1000 sample trajectories, with standard deviation bands shown in cyan. The initial condition is a vector of magnitude $|x(0)| = 5$ and random direction. The inset shows the end of the simulation in greater detail. (b) Evolution of the mean $\mathbb{E}(\tau_i)$ and (inset) histogram of τ_i. $mean(\tau_i) = 0.0104$ s, $std(\tau_i) = 7.54 \times 10^{-4}$ s, and $\min(\tau_i) = 3.78 \times 10^{-4}$ s.

the average $\mathbb{E}(\tau_i)$ of these inter-execution times over the 1000 simulations. For comparison, the inter-execution times in [248] range from 0.0058 s to 0.0237 s. As the samples approach the origin, the average inter-execution times $\mathbb{E}(\tau_i)$ increase (Fig. 7.1(b)), a trend that can be seen in previous deterministic works [18, 248]. With respect to our triggering condition, this is because, for larger $|x_i|$, $A(|x_i|,\tau_i)$ and $C(|x_i|,\tau_i)$ increase faster with τ_i, and $B(|x_i|,\tau_i)$ and $D(|x_i|,\tau_i)$ will decrease faster. These quantities appear on opposing sides of an inequality in (7.39) and (7.45), and consequently, for certain values of θ, these relations will hold for shorter periods of time with larger $|x_i|$.

In the case where $\sigma = 0$ and $\sigma_x = 0.1$, we can obtain several possible execution rules. The first is (7.50) with $d_\alpha = 1$, but this only guarantees practical stability in the mean square sense. Since $g(0,0) = 0$, we can also use the execution rule

$$C(|x_i|,\tau_i) \leq \theta D(|x_i|,\tau_i) + d_\alpha, \tag{7.51}$$

which also guarantees practical p-moment stability. However, for this system, we can also set $d_\alpha = 0$ in order to guarantee p-moment stability according to

$$\begin{aligned}\mathbb{E}\mathcal{L}V &\leq -(a-\sigma_x^2)\mathbb{E}(|x|^2) + b\sqrt{\mathbb{E}(|e|^2)}\sqrt{\mathbb{E}(|x|^2)} \\ &\leq -(a-\sigma_x^2 - b\sqrt{\theta})\mathbb{E}(|x|^2)\end{aligned}$$

Examining the form of $C(|x_i|,\tau_i)$ and $D(|x_i|,\tau_i)$ in (7.51), note that the triggering rule can be made independent of x_i when $d_\alpha = 0$, and, therefore, it reduces to a periodic update rule. In this case, the inter-execution time τ_i can be found as

$$\tau_i = \frac{1}{4\sqrt{L_m}+3L_m}\ln\left(\frac{(1-\theta)^2+\sqrt{(1-\theta)^2+2\theta\frac{6+5\sqrt{L_m}}{1+\sqrt{L_m}}}}{2}\right)$$

Since the periodic update rule may result in shorter inter-execution times, we first apply (7.51) with $d_\alpha > 0$ (a self-triggered approach) before switching to the periodic rule to ensure asymptotic stability. Fig. 7.2 shows $\mathbb{E}(|x|^2)$ and $\mathbb{E}(\tau_i)$ based on 1000 simulations using each of these three update rules (with $A(\cdot,\cdot)$, $B(\cdot,\cdot)$, and d_α; $C(\cdot,\cdot)$, $D(\cdot,\cdot)$, and d_α; and the periodic update based on $C(\cdot,\cdot)$ and $D(\cdot,\cdot)$). The second rule (7.51) results in the longest inter-execution times on average, and it does not appear necessary to switch to the periodic rule in order to achieve asymptotic stability in this example.

7.5.2 Stochastic Nonlinear Wheeled Cart System

This example is based on the wheeled cart control problem from [216, Equation (54)]. It describes a wheeled cart that should steer a point on its body to the origin (see Fig. 7.3(a)) with free final heading angle ϕ. We have added a stochastic disturbance

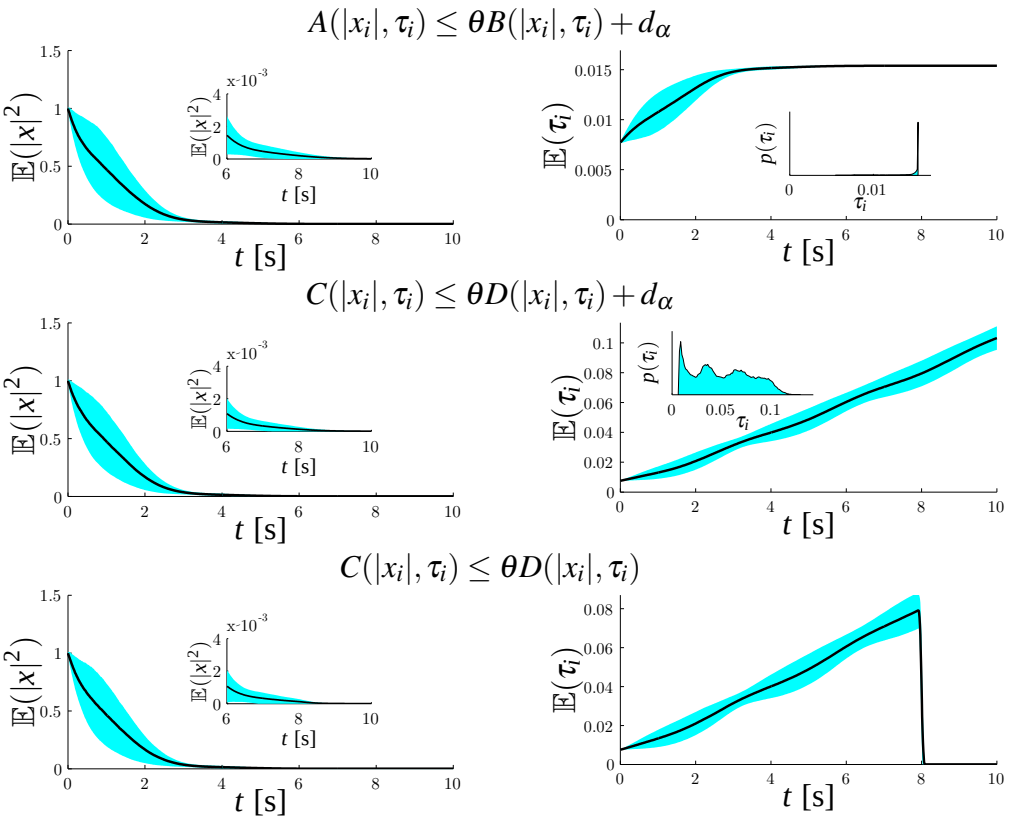


Figure 7.2: Linear system from [248] using three different sampling strategies. First row: With $A(|x_i|,\tau_i) \leq \theta B(|x_i|,\tau_i) + d_\alpha$, evolution of $\mathbb{E}(|x|^2)$ and $\mathbb{E}(\tau_i)$ for over a 10 s simulation, averaged over 1000 sample trajectories, and with an initial condition of magnitude $|x(0)| = 1$. $mean(\tau_i) = 0.0143$ s, $std(\tau_i) = 6.2 \times 10^{-4}$ s, $\min(\tau_i) = 0.0054$ s. Second row: The same quantities using $C(|x_i|,\tau_i) \leq \theta D(|x_i|,\tau_i) + d_\alpha$. $mean(\tau_i) = 0.0511$ s, $std(\tau_i) = 0.002$ s, $\min(\tau_i) = 0.0056$ s. Third row: Initially, the same as the second row, the update rule is changed to a periodic condition ($\tau_i = 3.5 \times 10^{-5}$ s) at $t = 8$ s.

with constant intensity to this example, which may describe motion in an uncertain environment. The system is described by the equation

$$\mathrm{d}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & x_2 \\ 0 & -(0.1+x_1) \end{bmatrix} u\mathrm{d}t + \sigma \begin{bmatrix} \mathrm{d}w_1 \\ \mathrm{d}w_2 \end{bmatrix} \tag{7.52}$$

where the control $u = [v,\ \dot{\phi}]^\mathsf{T}$ (see Fig. 7.3(a)). The feedback law $u = [k_1 x_1, k_2 x_2]^\mathsf{T}$ with positive gains k_1 and k_2 [216] can be shown to asymptotically stabilize $[x_1, x_2]^\mathsf{T}$ to a

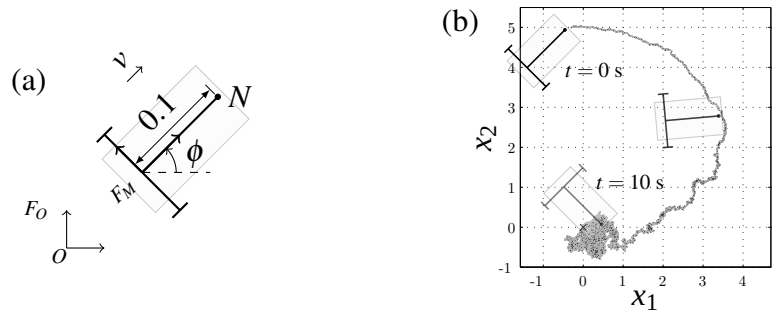


Figure 7.3: (a) Diagram of wheeled cart that should drive a point N that is a distance of 0.1 from its wheel axis to the origin O. The coordinates of $\overrightarrow{NO}$ in the mobile frame F_M are $[x_1, x_2]^\intercal$. See [216] for details. (b) An example trajectory under a self-triggered implementation (cart drawn not to scale and with arbitrary heading angle).

disc of radius $\sigma/\sqrt{k_1 \wedge 0.1k_2}$. Considering now a sampled-data implementation with $u_1 = k_1(x_1 + e_1)$ and $u_2 = k_2(x_2 + e_2)$,

$$\mathrm{d}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & x_2 \\ 0 & -(0.1 + x_1) \end{bmatrix} \begin{bmatrix} k_1(x_1 + e_1) \\ k_2(x_2 + e_2) \end{bmatrix} \mathrm{d}t + \sigma \begin{bmatrix} \mathrm{d}w_1 \\ \mathrm{d}w_2 \end{bmatrix} \tag{7.53}$$

With the choice of $V(x) = \frac{1}{2}|x|^2$ as a Lyapunov function,

$$\begin{aligned} \mathcal{L}V &= -k_1 x_1^2 - 0.1k_2 x_2^2 - k_1 x_1 e_1 - 0.1k_2 x_2 e_2 + \sigma^2 \\ &\leq -\underline{k}|x|^2 + \overline{k}|x||e| + \sigma^2 \end{aligned}$$

where $\underline{k} = k_1 \wedge 0.1k_2$ and $\overline{k} = k_1 \vee 0.1k_2$. Taking the expectation of both sides and applying Hölder's inequality,

$$\begin{aligned} \mathbb{E}\mathcal{L}V &\leq -\underline{k}\mathbb{E}(|x|^2) + \overline{k}\mathbb{E}(|x||e|) + \sigma^2 \\ &\leq -\underline{k}\mathbb{E}(|x|^2) + \overline{k}\sqrt{\mathbb{E}(|x|^2)}\sqrt{\mathbb{E}(|e|^2)} + \sigma^2 \\ &\leq -\underline{k}\mathbb{E}(|x|^2) + \overline{k}\sqrt{\mathbb{E}(|x|^2) + \theta_d\sigma^2}\sqrt{\mathbb{E}(|e|^2)} + \sigma^2 \end{aligned}$$

If we were to assume that

$$\mathbb{E}(|e|^2) \leq \theta \mathbb{E}(|x|^2) + \theta_d^2 \sigma^2 \tag{7.54}$$

for some $0 < \theta < (\underline{k}/\overline{k})^2$, then

$$\mathbb{E}\mathcal{L}V \leq -(\underline{k} - \overline{k}\sqrt{\theta})\mathbb{E}(|x|^2) + \sigma^2(1 + \frac{\theta_d}{\sqrt{\theta}}\overline{k}).$$

Using Gronwall's inequality, one can then show that the triggering rule (7.54) will steer the cart's point $[x_1, x_2]^\intercal$ to the origin such that for a sufficiently large t,

$$\mathbb{E}(|x(t)|) \leq \mathbb{E}(|x(t)|^2) \leq \sigma^2 \frac{1 + \theta_d \overline{k}/\sqrt{\theta}}{\underline{k} - \overline{k}\sqrt{\theta}}.$$

Similarly to the previous example, we set $\alpha_v(|x|) = |x|$, and $\lambda_c(|e|) = |e|$ and use the triggering rule (7.54) with $d_\alpha = \theta_d^2 \sigma^2 = 1$. For simulation, we let $k_1 = k_2 = 0.5$, which requires that $0 < \theta < 0.01$, and so we choose $\theta = 0.009$. With $\sigma = 0.4$, the monotone growth coefficient is $K = \overline{k}/2 \vee \sigma^2 = 0.25$. An example trajectory for the resulting sampled-data scheme can be found in Fig. 7.3(b). Fig. 7.4 shows the mean square $E(|x(t)|^2)$ of 1000 trajectories of $x(t)$ starting from a random vector of magnitude $|x(0)| = 5$. As with the previous example, the mean inter-execution times $\mathbb{E}(\tau_i)$ increase as $|x_i|$ approaches the origin (Fig. 7.4(b)).

7.6 Conclusions

This paper presents a self-triggered control scheme for state-feedback controlled stochastic differential equations. Since the inequality-based sampling conditions found in previous event- and self-triggered control works may be instantaneously violated in

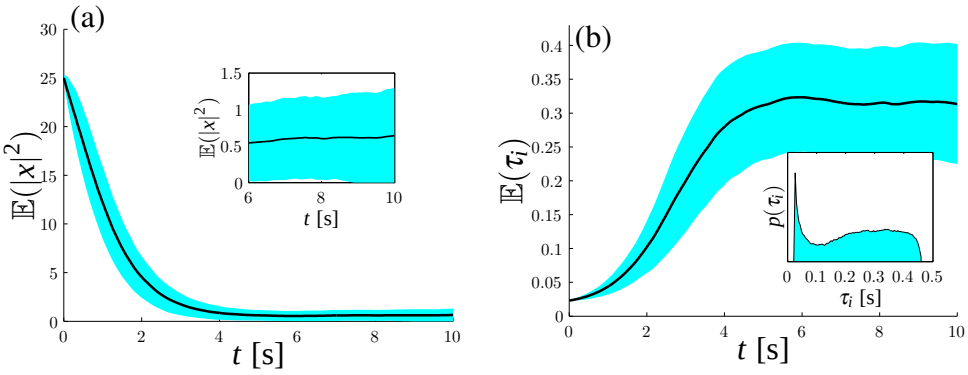


Figure 7.4: Nonlinear wheeled cart system from [216] with stochasticity added. (a) Evolution of $\mathbb{E}(|x|^2)$ for 1000 simulations with an initial condition of a random vector of magnitude $|x(0)| = 5$, and with standard deviation bands shown in cyan. The inset shows the end of the simulation in greater detail. (b) Evolution of the mean $\mathbb{E}(\tau_i)$ over the 1000 simulations with standard deviation bands shown. The inset shows a histogram of these times. $mean(\tau_i) = 0.2375$ s, $std(\tau_i) = 0.0282$ s, and $\min(\tau_i) = 0.195$ s.

the presence of the stochastic noise considered in this paper, we instead focus on the statistics of the state distribution. These quantities can be predicted based on the last-observed state and are used here to develop a self-triggered control scheme. Since for many systems there is no guarantee that the stochasticity will diminish at the origin, we have considered alongside the error due to sampling a second disturbance caused by non-vanishing noise. We presented two triggering conditions based on whether or not this noise vanishes at the origin. The scheme is shown to produce strictly positive inter-execution times that guarantee (practical) p-moment stability of the process.

In future work, elongation of the task periods may be obtainable by taking into account the direction of the error as compared to the current state instead of just its magnitude. Further improvements may be possible by treating the error due to sampling as a delay and more directly applying the stability result from stochastic differential delay systems that is used in this work. The robustness of our scheme to a delay between

state sampling and implementation of the updated control, i.e., a task delay, will be also examined in future work, as well as the application of this scheme for stochastic problems where control updates or state sampling are expensive or limited, e.g., multi-agent robotic systems.

Chapter 8

Concluding Remarks and Future Work

While each of the preceding paper prepints were independently concluded, we offer here some overall remarks on this dissertation and directions for future work.

This dissertation addressed a series of related problems in the guidance and navigation of autonomous vehicles in the presence of uncertainty. With primary focus on the Dubins vehicle and the unicycle, we are motivated by the need to develop control policies for the vehicles that a) achieve a certain level of performance with respect to individual or group objectives; b) abide by kinodynamic constraints; c) are as transparent as possible; and d) anticipate the intrinsically stochastic nature of the vehicles, their control algorithms, and the environment.

We began with a new take on a frequently-studied problem in the UAV community — the feedback control of a Dubins vehicle. At first, we tackled a tracking problem for the nonholonomic vehicle to maintain a nominal standoff distance from a target with unknown future motion. Unlike previous works, we explicitly modeled the stochasticity

of the target by assuming it could be robustly described by Brownian motion. The possibility for the loss of target observations was included by considering stochastic transitions in and out of the act of observing the target. An optimal feedback control, computed from the HJB equation, anticipated the future motion of this target, even if its trajectory was deterministic. Similar to this problem was the navigation of a Dubins vehicle in a stochastic wind. Models for the stochastic variation of the wind were used in the computation of an optimal feedback control policy to steer the Dubins vehicle to a target in minimum time. In this work, we also characterized how the optimal feedback control policy changes in the presence of stochasticity.

For a control designer, the choice of a turning rate in the deterministic versions of these vehicle control problems seems quite intuitive. For example, for the Dubins vehicle in wind problem, it seems reasonable that a Dubins vehicle should attempt to steer its velocity vector toward its target, and, as we have seen, this turns out to be the optimal turning rate in the deterministic case. A clear advantage to this simplicity is that an analytical feedback control can be developed. We have also seen that the optimal feedback control in the stochastic versions is not always as intuitive. For example, while in some instances, the Dubins vehicle should point its heading toward the target, there are also cases when the Dubins vehicle should wager whether or not it can hit its target, at risk of being blown off course. Or, with greater noise intensities, it should ignore this possibility and again revert to pointing its velocity vector toward the target. The noise therefore adds an element to the problem not seen in the deterministic versions,

and unfortunately, there is not yet an obvious or clear path to writing down a feedback control law that is designed to perform well with respect to the modeled noise. In other words, the role of noise in designing feedback policies is not well understood.

Granted, there is a rich theory of stochastic stability and stabilization that can be used to develop a stabilizing control policy. However, these methods focus on whether or not a certain function of state variables is decreasing in time, and they do not necessarily follow from human intuition about the effect of noise on a problem. This can be contrasted against the deterministic case, in which stabilization methods often do permit some degree of qualitative behavior design when creating a feedback control. Consider the formation control problem, for example. A feedback control that steers each vehicle's heading angle toward the minimum of some artificial potential function, for example, might seem like a reasonable first step. Use of Lyapunov function techniques can then be used to show that this control law asymptotically leads to a formation. How, then, do we account for the stochasticity of the problem?

This led us to the next topic tackled in this work, which proposed an enhancement to the deterministic, analytic, and intuitive feedback control laws with a numerically-computed additive input. The additive correction terms were computed from the HJB equation, but, due to the large dimension of the state space, this could not be performed on a discrete state space. Instead, using a path integral representation, we developed an equivalent representation on a continuous state space and showed how a Kalman smoothing algorithm can be used to compute the corrective control in real-time. The

resulting feedback controls were optimal and robust to the uncertainties introduced by the vehicles or environment. In addition to being used in real-time for vehicle control, we showed how the speed of the Kalman smoothing algorithm can be useful for analyzing the deterministic feedback controls for the purpose of developing analytical improvements. We envision that the use of this approach may aide in elucidating some of the mystery in analytic stochastic feedback control design.

Finally, we have examined a key issue in the implementation of control laws on board the autonomous vehicle systems. In particular, we developed a scheme in which the discrete times when a vehicle control loop should be updated are predicted using the last-observed state (a self-triggered approach).

The four preprints included in the previous chapters represent milestones in the line of work leading to this dissertation. Continuing on this path, there are many possible directions for future work that build upon our results, and we now highlight some of them.

A more thorough analysis of the bifurcations in the control law based on the parameters (e.g., noise intensity) would be an interesting direction for future research. A quantitative characterization of the qualitative anticipatory behavior of the vehicles would be revealing in light of the puzzling role of noise in feedback control design. For example, the location of the boundary in state space at which the (minimum-time) Dubins vehicle should no longer steer directly toward its target moves with increasing noise, but later disappears with an even greater noise intensity. Additionally, as pre-

viously mentioned, our numerical scheme for value iterations fails to converge on an optimal policy for some combinations of parameters. This could be due the fact that, for some levels of stochasticity, there is no admissible control input with respect to the objective at hand. The point at which a target becomes "too" stochastic could be examined as part of this bifurcation analysis. Characterizing the change in the control law due to noise intensity would also help with the problem of scaling a control law for use in different parameter regimes, as discussed in previous sections.

One of the themes discussed in this work is the effect on outdated samples on the feedback control loop. At first, this was an outdated sample due to the loss of target observations in the Dubins tracking problem. In this case, a probability distribution of where the target might be found was used to maintain a good tracking performance during target loss. However, we did not explicitly examine the stability of the control loop under observation loss, and, in fact, there were rates of target loss for which value iteration did not converge to an optimal policy, suggesting that the duration of target observation loss in these instances was too long to maintain stability. The next discussion of outdated samples appeared in the sample-and-hold scheme of the self-triggering problem. Here, we were concerned with maintaining stability under the scheme for some duration, but not with how to keep a good performance. There is some overlap between these two problems, and they appear to be tackling the same issue from different angles. A possible avenue of future work, then, could examine the steps necessary to initially achieve a good performance during a sample hold, and, as the length of time since the

last sample increase, if that becomes difficult, then revert purely to stabilization.

Another application of the outdated samples not explored in this dissertation, but closely related to the presented topics, is the option for outdated samples to be used in lieu of observations of some proximate neighbors in large-scale control problems. Along these lines, a vehicle could apply a control that is optimal with respect to the current state of one proximate vehicle, but simply non-destabilizing with respect to outdated samples of other neighbors. Based on predictions of the remaining vehicles, it could then switch focus among the vehicles. This would decrease the size of the state space used to compute the optimal control and would also reduce the overhead associated with observations of all proximate vehicles at once, and is worthy of future consideration.

The Kalman smoother-based algorithms presented in this work were shown to be effective in creating a formation among the vehicles (even in the case where the underlying deterministic flocking control is removed), but we have not presented a proof of convergence. In fact, this problem can be interpreted as a type of distributed estimation problem that we have not seen elsewhere. In our algorithm, each vehicle estimates its trajectory and those of its neighbors over some planning horizon. When a vehicle executes the first step of this control and subsequently observes the state of its neighbors, as detailed in Chapter 6, the vehicles are essentially exchanging partial information about their estimates, i.e., they only share the first step of their estimated trajectory. As the process repeats, the formation is reached if this distributed estimation problem also

reaches a consensus. Developing the conditions under which this estimation problem converges would, threfore, be a pertinent direction for research.

The implementation of the controllers presented in this dissertation on real vehicles is likely to present additional problems that we have not yet tackled. Some of the potential implementation issues were detailed in Section 3.4 and Chapter 7. However, our noise-anticipating control design will hopefully mitigate some of the potential hurdles that can arise in implementation. Experiments with real robots are underway.

Bibliography

[1] K. Årzén. A simple event-based PID controller. In *Proceedings of the 14th IFAC World Congress*, Beijing, China, 1999.

[2] P. Abbeel and K. Goldberg. LQG-MP : Optimized Path Planning for Robots with Motion Uncertainty and Imperfect State Information. *International Journal of Robotics Research*, 30(7):895–913, 2010.

[3] W. Aggoune, B. Castillo, and S. Di Gennaro. Self-triggered robust control of nonlinear stochastic systems. *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 7547–7552, Dec. 2012.

[4] N. Aghasadeghi and T. Bretl. Maximum Entropy Inverse Reinforcement Learning in Continuous State Spaces with Path Integrals. In *Proceedings of the 2011 IEEE Conference on Intelligent Robots and Systems*, pages 1561–1566, 2011.

[5] M. Aicardi, G. Casalino, A. Bicchi, and A. Balestrino. Closed loop steering of unicycle like vehicles via Lyapunov techniques. *IEEE Robotics & Automation Magazine*, (March):27–35, 1995.

[6] I. F. Akyildiz, D. Pompili, and T. Melodia. Underwater acoustic sensor networks: research challenges. *Ad Hoc Networks*, 3(3):257–279, May 2005.

[7] K. Alfriend, S. R. Vadali, P. Gurfil, J. How, and L. Breger. *Spacecraft Formation Flying: Dynamics, control, and navigation*. Butterworth-Heinemann, 2009.

[8] B. Anderson, B. Fidan, C. Yu, and D. Walle. UAV formation control: theory and application. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, pages 15–34. Springer-Verlag, 2008.

[9] R. Anderson, G. Dinolov, D. Milutinovic, and A. Moore. Maximally-informative ocean modeling system (ROMS) navigation of an AUV in uncertain ocean currents. In *ASME Dynamic Systems and Control Conference*, Fort Lauderdale, FL, 2012.

[10] R. P. Anderson and D. Milutinović. Dubins vehicle tracking of a target with unpredictable trajectory. In *in Proc. 4th ASME Dynamic Systems and Control Conference*, Arlington, VA, Oct. 31–Nov. 2, 2011.

[11] R. P. Anderson and D. Milutinović. A Stochastic approach to Dubins feedback control for target tracking. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3917–3922, San Francisco, CA, Sep. 25–30, 2011.

[12] R. P. Anderson and D. Milutinović. A Stochastic Optimal Enhancement of Feedback Control for Unicycle Formations. In *Proceedings of the International Symposium on Distributed Autonomous Robotic Systems*, Baltimore, MD, 2012.

[13] R. P. Anderson, D. Milutinović, and D. .V. Dimarogonas. Self-Triggered stabilization of continuous stochastic state-feedback controlled systems. In *Proceedings of the 2013 European Control Conference*, Zurich, Switzerland, 2013. IEEE.

[14] R. P. Anderson and D. Milutinovic. Distributed path integral feedback control based on Kalman smoothing for unicycle formations. In *Proceedings of the 2013 American Control Conference*, Washington, DC, 2013. IEEE.

[15] R. P. Anderson and D. Milutinović. Anticipating stochastic observation loss during optimal target tracking by a small aerial vehicle. In *Proceedings of the International Conference on Unmanned Aircraft Systems (ICUAS)*, pages 278–287, Atlanta, GA, 2013. IEEE.

[16] R. P. Anderson, E. Bakolas, D. Milutinović, and P. Tsiotras. Optimal Feedback Guidance of a Small Aerial Vehicle in the Presence of Stochastic Wind. *AIAA Journal of Guidance, Navigation, and Control*, 36(4):975–985, 2013.

[17] A. Anta and P. Tabuada. Self-triggered stabilization of homogeneous control systems. In *Proceedings of the 2008 American Control Conference*, pages 4129–4134, Seattle, WA, June 2008.

[18] A. Anta and P. Tabuada. To Sample or not to Sample: Self-Triggered Control for Nonlinear Systems. *IEEE Transactions on Automatic Control*, 55(9):2030–2042, Sept. 2010.

[19] A. Anta, P. Tabuada, and S. Member. Exploiting Isochrony in Self-Triggered Control. *IEEE Transactions on Automatic Control*, 57(4):950–962, 2012.

[20] D. Assaf, L. Shepp, and Y.-C. Yao. Optimal Tracking of Brownian Motion. *Stochastics and Stochastic Reports*, 54:233–245, 1995.

[21] K. Åström and B. Bernhardsson. Comparison of Riemann and Lebesgue sampling for first order stochastic systems. In *Proceedings of the 41st IEEE Conference on Decision and Control*, pages 2011–2016, Las Vegas, NV, 2002.

[22] M. Athans. The role and use of the stochastic linear-quadratic-Gaussian problem in control system design. *IEEE Transactions on Automatic Control*, 16(6):529–552, Dec. 1971.

[23] M. Athans and P. L. Falb. Time-Optimal Control for a Class of Nonlinear Systems. *IEEE Transactions on Automatic Control*, 8(4):1963, 1963.

[24] E. Bakolas. *Optimal steering for kinematic vehicles with applications to spatially distributed agents*. Ph.D. dissertation, School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, 2011.

[25] E. Bakolas and P. Tsiotras. Minimum-time paths for a small aircraft in the presence of regionally-varying strong winds. *AIAA Infotech@Aerospace, Atlanta, GA*, 2010.

[26] E. Bakolas and P. Tsiotras. Time-optimal synthesis for the Zermelo-Markov-Dubins problem: the constant wind case. In *Proceedings of American Control Conference*, pages 6163–6168, Baltimore, MD, June 30–July 2, 2010.

[27] E. Bakolas and P. Tsiotras. Feedback navigation in an Uncertain Flow Field and Connections with Pursuit Strategies. *Journal of Guidance, Control, and Dynamics*, 2011.

[28] E. Bakolas and P. Tsiotras. Optimal synthesis of the asymmetric sinistral/dextral Markov-Dubins problem. *Journal of Optimization Theory and Applications*, 150 (2):233–250, 2011.

[29] E. Bakolas and P. Tsiotras. Optimal Synthesis of the Zermelo - Markov - Dubins Problem in a Constant Drift Field. *Journal of Optimization Theory and Applications*, 2012.

[30] E. Bakolas and P. Tsiotras. Feedback navigation in an uncertain flow-field and connections with pursuit strategies. *Journal of Guidance, Control, and Dynamics*, 35(4):1268–1279, 2012.

[31] T. Balch and R. C. Arkin. Communication in Reactive Multiagent Robotic Systems. *Autonomous Robots*, 1:27–52, 1994.

[32] D. Balkcom and M. Mason. Time optimal trajectories for bounded velocity differential drive vehicles. *The International Journal of Robotics Research*, 21

(3):199-218, 2002.

[33] D. J. Balkcom, P. A. Kavathekar, and M. T. Mason. The time-optimal trajectories for an omni-directional vehicle. *The International Journal of Robotics Research*, 25(10):985-999, 2006.

[34] A. Balluchi, A. Bicchi, A. Balestrino, and G. Casalino. Path tracking control for Dubin's cars. In *Proceedings of the 1996 IEEE Conference on Robotics and Automation*, volume 4, pages 3123–3128. IEEE, 1996.

[35] A. Balluchi, A. Bicchi, B. Piccoli, and P. Soueres. Stability and robustness of optimal synthesis for route tracking by Dubins' vehicles. *Proceedings of the 39th IEEE Conference on Decision and Control*, pages 581–586, 2010.

[36] M. Bardi and I. Capuzzo-Dolcetta. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Birkhauser Boston, 1997.

[37] J. Beaman. Non-linear quadratic gaussian control. *International Journal of Control*, 39(2):343–362, 1984.

[38] R. Beard, T. Mclain, D. Nelson, D. Kingston, and D. Johanson. Decentralized Cooperative Aerial Surveillance Using Fixed-Wing Miniature UAVs. *Proceedings of the IEEE*, 94(7):1306–1324, July 2006.

[39] R. W. Beard and T. W. McLain. *Small Unmanned Aircraft: Theory and Practice*. Princeton University Press, Princeton, NJ, 2012.

[40] J. G. Bellingham and K. Rajan. Robotics in remote and hostile environments. *Science (New York, N.Y.)*, 318(5853):1098–102, Nov. 2007.

[41] R. Bellman. *Adaptive Control Processes: A Guided Tour*. Princeton University Press, Princeton, 1961.

[42] M. D. Benedetto, M. Domenica, S. Di Gennaro, and A. D'Innocenzo. Digital self-triggered robust control of nonlinear systems. *Journal of Control*, (in press) 2013.

[43] V. Beneš, L. A. Shepp, and H. S. Witsenhausen. Some solvable stochastic control problems. *Stochastics*, 4:39–83, 1980.

[44] D. P. Bertsekas. Dynamic Programming and Suboptimal Control: A Survey from ADP to MPC. *European Journal of Control*, 11(4-5):310–334, Jan. 2005.

[45] D. P. Bertsekas. *Dynamic Programming and Optimal Control, Volume II: Ap-*

proximate Dynamic Programming. 4th edition, 2012.

[46] D. P. Bertsekas and J. N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, Belmont, MA, 1996. Chap. 2, pp. 12–57.

[47] D. P. Bertsekas and J. N. Tsitsiklis. Comments on "Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules". *IEEE Transactions on Automatic Control*, 52(5):968–969, 2007.

[48] A. Bloch. *Nonholonomic mechanics and control*, volume 24. Springer, 2003.

[49] J. D. Boissonnat and X. N. Bui. Accessibility region for a car that only moves forward along optimal paths. Research Note 2181, Institut National de Recherche en Informatique et en Automatique, Sophia-Antipolis, France, 1994.

[50] J.-D. Boissonnat, A. Cerezo, and J. Leblond. Shortest paths of bounded curvature in the plane. *Journal of Intelligent & Robotic Systems*, 11(1-2):5–20, Mar. 1994.

[51] A. N. Borodin and P. Salminen. *Handbook of Brownian Motion: Facts and Formulae*. Birkhauser, Basel, Switzerland, 2nd edition, 2002.

[52] F. Borrelli, T. Keviczky, and G. Balas. Collision-free UAV formation flight using decentralized optimization and invariant sets. *2004 43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No.04CH37601)*, pages 1099–1104 Vol.1, 2004.

[53] B. Bouilly, T. Simeon, and R. Alami. A numerical technique for planning motion strategies of a mobile robot in presence of uncertainty. In *Proceedings of the 1995 IEEE International Conference on Robotics and Automation*, pages 1327–1332, Nagoya, Japan, 1995. IEEE.

[54] A. Bressan, Noncooperative Differential Games. *Milan J. Math.* 79(2011):357–427, 2011.

[55] A. Brezoescu, R. Lozano, T. Espinoza, and P. Castillo. Adaptive trajectory following for a xed-wing UAV in presence of crosswind. 4(1):257–271, 2013.

[56] X. Bui, J. Boissonnat, P. Soueres, and J.-P. Laumond. Shortest path synthesis for Dubins non-holonomic robot. In *Proceedings of the 1994 IEEE Conference on Robotics and Automation*, pages 2–7, 1994.

[57] F. Bullo, J. Cortes, and S. Martinez. *Distributed control of robotic networks: A mathematical approach to motion coordination algorithms*. Princeton University Press, Princeton, NJ, 2009.

[58] Y. Cao and A. Fukunaga. Cooperative mobile robotics: Antecedents and directions. *Proceedings of the 1995 IEEE/RSJ Conference on Human Robot Interaction and Cooperative Robots*, 1995.

[59] Y. Cao, W. Yu, W. Ren, and G. Chen. An Overview of Recent Progress in the Study of Distributed Multi-Agent Coordination. *IEEE Transactions on Industrial Informatics*, 9(1):427–438, Feb. 2013.

[60] N. Ceccarelli, J. J. Enright, E. Frazzoli, S. J. Rasmussen, and C. J. Schumacher. Micro UAV Path Planning for Reconnaissance in Wind. *2007 American Control Conference*, pages 5310–5315, July 2007.

[61] N. Ceccarelli, J. J. Enright, E. Frazzoli, S. J. Rasmussen, and C. J. Schumacher. Micro UAV path planning for reconnaisance in wind. In *Proceedings of American Control Conference*, pages 5310–5315, New York City, NY, July 11–13, 2007.

[62] Z. Chao, S.-L. Zhou, L. Ming, and W.-G. Zhang. UAV Formation Flight Based on Nonlinear Model Predictive Control. *Mathematical Problems in Engineering*, 2012:1–15, 2012.

[63] C. Charalambous. The role of information state and adjoint in relating nonlinear output feedback risk-sensitive control and dynamic games. *IEEE Transactions on Automatic Control*, 42(8):1163–1170, 1997.

[64] H. Chitsaz and S. M. LaValle. Time-optimal paths for a Dubins airplane. In *Proc. 46th IEEE Conference on Decision and Control*, pages 2379–2384. IEEE, 2008.

[65] S.-o. L. Y.-j. Cho and M. H.-b. B.-j. You. A Stable Target-Tracking Control for Unicycle Mobile Robots. In *Proceedings 2000 IEEE Conference on Intelligent Robots and Systems (IROS)*, pages 1822–1827. IEEE, 2000.

[66] E. A. Codling, M. J. Plank, and S. Benhamou. Random walk models in biology. *Journal of the Royal Society Interface*, 5(25):813–34, Aug. 2008.

[67] R. Cortez, H. Tanner, and R. Lumia. Distributed Robotic Radiation Mapping. In *Experimental Robotics*, pages 147–156. Springer, 2009.

[68] E. Cristiani and P. Martinon. Initialization of the Shooting Method via the Hamilton-Jacobi-Bellman Approach. *Journal of Optimization Theory and Applications*, 146(2):321–346, Feb. 2010.

[69] A. Czirók, A-L. Barabási, and T. Vicsek. Collective motion of Self-Propelled

Particles: Kinetic Phase Transition in One Dimension. *Physical Review Letters*, 82(1):209–212, 1999.

[70] R. E. Davis, N. E. Leonard, and D. M. Fratantoni. Routing strategies for underwater gliders. *Deep Sea Research Part II: Topical Studies in Oceanography*, 56 (3-5):173–187, Feb. 2009.

[71] A. De Luca and G. Oriolo. Modeling and Control of Nonholonomic Mechanical Systems. In J. Angeles and A. Kecskemethy, editors, *Kinematics and Dynamics of Multibody Systems*, chapter 7. Springer-Verlag, 1995.

[72] J. Desai, J. Ostrowski, and V. Kumar. Modeling and control of formations of nonholonomic mobile robots. *IEEE Transactions on Robotics and Automation*, 17(6):905–908, 2001.

[73] D. Dimarogonas. A connection between formation control and flocking behavior in nonholonomic multiagent systems. *International Conference on Robotics and Automation*, 2006.

[74] D. Dimarogonas. On the rendezvous problem for multiple nonholonomic agents. *IEEE Transactions on Automatic Control*, 52(5):916–922, 2007.

[75] D. Dimarogonas, E. Frazzoli, and K. Johansson. Distributed event-triggered control strategies for multi-agent systems. *IEEE Transactions on Automatic Control*, 57(5):1291–1297, 2012.

[76] D. V. Dimarogonas and E. Frazzoli. Analysis of decentralized potential field based multi-agent navigation via primal-dual Lyapunov theory. *49th IEEE Conference on Decision and Control (CDC)*, (1):1215–1220, Dec. 2010.

[77] D. V. Dimarogonas, E. Frazzoli, and K. H. Johansson. Distributed self-triggered control for multi-agent systems. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 6716–6721, Atlanta, GA, Dec. 2010.

[78] X. C. Ding, A. R. Rahmani, and M. Egerstedt. Multi-UAV convoy protection: an optimal approach to path planning and coordination. *IEEE Transactions on Robotics*, 26(2), Apr. 2010.

[79] V. Dobrokhodov, K. Jones, and R. Ghabcheloo. Vision-Based Tracking and Motion Estimation for Moving targets using Small UAVs. *2006 American Control Conference*, pages 1428–1433, 2006.

[80] I. Dolinskaya. *Optimal path finding in direction, location, and time dependent environments*. PhD thesis, The University of Michigan, 2009.

[81] P. V. Dooreen. Gramian Based Model Reduction of Large-Scale Dynamical Systems. 1999.

[82] L. E. Dubins. On curves of minimal length with a constraint on average curvature, and with prescribed initial and terminal positions and tangents. *American Journal of Mathematics*, 79(3):497–516, 1957.

[83] V. Duindam, J. Xu, R. Alterovitz, S. Sastry, and K. Goldberg. 3D Motion Planning Algorithms for Steerable Needles Using Inverse Kinematics. *The International journal of robotics research*, 57:535–549, Jan. 2009.

[84] W. B. Dunbar and R. M. Murray. Distributed Receding Horizon Control with Application to Multi-Vehicle Formation Stabilization. *Automatica*, pages 1–47, 2004.

[85] K. Dvijotham and E. Todorov, Linearly Solvable Optimal Control. In *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*, chapter 6. Wiley and IEEE Press, 2012.

[86] M. Egerstedt, X. Hu, and a. Stotsky. Control of mobile platforms using a virtual vehicle approach. *IEEE Transactions on Automatic Control*, 46(11):1777–1782, 2001.

[87] G. Elkaim and R. Kelbley. A Lightweight Formation Control Methodology for a Swarm of Non-Holonomic Vehicles. In *IEEE Aerospace Conference*, Big Sky, MT, 2006. IEEE.

[88] J. Enright, E. Frazzoli, K. Savla, and F. Bullo. On Multiple UAV Routing with Stochastic Targets: Performance Bounds and Algorithms. In *AIAA Guidance, Navigation, and Control Conference*, San Francisco, CA, Aug. 2005.

[89] J. J. Enright and E. Frazzoli. UAV routing in a stochastic, time-varying environment. In *Proceedings of the 16th IFAC World Congress*, Czech Republic, 2005.

[90] J. J. Enright, E. Frazzoli, K. Savla, and F. Bullo. On multiple UAV routing with stochastic targets: performance bounds and algorithms. In *Proc. AIAA Guidance, Navigation, and Control Conference and Exhibit*, San Francisco, 2005.

[91] A. Eqtami, D. V. Dimarogonas, and K. J. Kyriakopoulos. Event-Based Model Predictive Control for the Cooperation of Distributed Agents. In *Proceedings of the 2012 American Control Conference*, Montreal, QC, pages 6473–6478, 2012.

[92] M.S. Fadali and A. Visioli *Digital Control Engineering: Analysis and Design*.

2nd ed., Academic Press, Waltham, MA, 2013.

[93] P. L. Falb, W. A. Wolovich, L. M. Silverman, W. M. Wonham, I. Sugiura, R. W. Shields, B. Pearson, T. Liu, and W. K. Chen. Determining Balancing and Stochastic Model Reduction. *IEEE Transactions on Automatic Control*, AC-30 (9):921–923, 1985.

[94] S. Fan, Z. Feng, and L. Lian. Collision free formation control for multiple autonomous underwater vehicles. In *Proceedings of the IEEE Oceans 2010*, pages 1–4, May 2010.

[95] J. A. Fax. *Optimal and Cooperative Control of Vehicle Formations*. Ph.D. dissertation, California Institute of Technology. Pasadena, CA, 2002.

[96] W. Fleming and H. Soner. Logarithmic Transformations and Risk Sensitivity. In *Controlled Markov Processes and Viscosity Solutions*, chapter 6. Springer, Berlin, 1993.

[97] W. Fleming and H. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer, 2nd edition, 2006.

[98] T. Fraichard and R. Mermond. Path Planning with uncertainty for Car-Like Robots. In *Proceedings of the 1998 IEEE International Conference on Robotics & Automation*, number May, pages 27–32, Leuven, Belgium, 1998. IEEE.

[99] R. Freeman and J. Primbs. Control Lyapunov functions: new ideas from an old source. In *Proceedings of 35th IEEE Conference on Decision and Control*, volume 4, pages 3926–3931, Kobe, Japan, 1996. IEEE.

[100] R. A. Freeman and P. V. Kokotovic. Inverse Optimality in Robust Stabilization. *SIAM Journal on Control and Optimization*, 34(4):1365, 1996.

[101] M. Freidlin. *Functional Integration and Partial Differential Equations*. Princeton University Press, Princeton, NJ, 1985.

[102] D. Galzi and Y. Shtessel. UAV formations control using high order sliding modes. in *Proceedings of the 2006 American Control Conference*, 2006.

[103] F. Gao and F. Yuan. Finite-time stabilization of stochastic nonholonomic systems and its application to mobile robot. pages 1–21.

[104] E. Garcia and P. J. Antsaklis. Model-Based Event-Triggered Control for Systems With Quantization and Time-Varying Network Delays. *IEEE Transactions on Automatic Control*, 58(2):422–434, Feb. 2013.

[105] C. Gardiner. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer, fourth edition, 2009. Chaps. 1–4., pp. 1–116.

[106] V. Gazi and B. Fidan. Coordination and Control of Multi-agent Dynamic Systems: Models and Approaches. In E. Sahin, W. M. Spears, and A. F. T. Winfield, editors, *Lecture Notes in Computer Science vol 4433: Swarm Robotics*, pages 71–102. Springer-Verlag, Berlin, 2007.

[107] V. Gazi and K. M. Passino. *Swarm Stability and optimization*. Springer, Berlin, 2011.

[108] A. Gelb. *Applied Optimal Estimation*. The M.I.T. Press, Cambridge, MA, 1974.

[109] J. Ghommam, H. Mehrjerdi, M. Saad, and F. Mnif. Formation path following control of unicycle-type mobile robots. *Robotics and Autonomous Systems*, 58 (5):727–736, May 2010.

[110] B. F. Giulietti, L. Pollini, and M. Innocenti. Autonomous Formation Flight. *IEEE Control Systems Magazine*, pages 34–44, December 2000.

[111] H. Goldstein. *Classical Mechanics*. Addison-Wesley, 2nd edition, 1980.

[112] F. B. Hanson. Techniques in Computational Stochastic Dynamic Programming. In *Stochastic Digital Control System Techniques*, vol. 76. pages 103–162, 1996.

[113] A. Hayes, A. Martinoli, and R. Goodman. Distributed odor source localization. *IEEE Sensors Journal*, 2(3):260–271, June 2002.

[114] S. Hota and D. Ghose. A modified Dubins method for optimal path planning of a miniature air vehicle converging to a straight line path. In *American Control Conference, 2009. ACC'09.*, pages 2397–2402. IEEE, 2009.

[115] J. Hu, G. Chen, and H.-x. Li. Distributed Event-triggered Tracking Control fo leader-Follower Multii–agent Systems with Communication Delay. *Kybernetika*, 47(4):630–643, 2011.

[116] L. Huang and X. Mao. On input to state stability of stochastic retarded systems with Markovian switching. *IEEE Transactions on Automatic Control*, 54(8): 1898–1902, 2009.

[117] S. Ikenoue, M. Asada, and K. Hosoda. Cooperative behavior acquisition by asynchronous policy renewal that enables simultaneous learning in multiagent environment. *IEEE/RSJ International Conference on Intelligent Robots and System*, (October):2728–2734, 2002.

[118] R. Isaacs. Games of pursuit. RAND Report P-257, RAND Corporation, Santa Monica, CA, 1951.

[119] S. Itani and M. a. Dahleh. On the Stochastic TSP for the Dubins Vehicle. In *Proceedings of the 2007 American Control Conference*, pages 443–448. Ieee, July 2007.

[120] R. Iyer, R. Arizpe, and P. Chandler. On the Existence and Uniqueness of Minimum Time Optimal Trajectory for a Micro Air Vehicle under Wind Conditions. *Emergent Problems in Nonlinear Systems and Control*, pages 57–77, 2009.

[121] a. Jadbabaie and a.S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, June 2003.

[122] A. Jadbabaie and J. Hauser. On the stability of unconstrained receding horizon control with a general terminal cost. In *Proceedings of the 40th IEEE Conference on Decision and Control*, volume 5, pages 4826–4831, Orlando, FL, 2001. IEEE.

[123] W. Just, H. Kantz, and C. R. Stochastic modelling: replacing fast degrees of freedom by noise. *Journal of Physics A: Mathematical and General*, 34(2001): 3199–3213, 2001.

[124] H. Kappen. Path integrals and symmetry breaking for optimal control theory. *Journal of statistical mechanics: theory and experiment*, 2005:P11011, 2005.

[125] H. Kappen. Linear Theory for Control of Nonlinear Stochastic Systems. *Physical Review Letters*, 95(20):1–4, Nov. 2005.

[126] H. Kappen, W. Wiegerinck, and B. van den Broek. A path integral approach to agent planning. *Autonomous Agents and Multi-Agent Systems*, 2007.

[127] H. J. Kappen, V. Gómez, and M. Opper. Optimal control as a graphical model inference problem. *Machine Learning*, 87(2):159–182, Feb. 2012.

[128] S. Karaman and E. Frazzoli. Sampling-based algorithms for optimal motion planning. *The International Journal of Robotics Research*, 2011.

[129] R. L. Kautz. Using chaos to generate white noise. *Journal of Applied Physics*, 86(10):5794, 1999.

[130] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, Upper Saddle River, New Jersey, 3rd edition, 2002.

[131] R. Khasminskii. *Stochastic Stability of Differential Equations*. Springer-Verlag, Berlin, 2nd edition, 2012.

[132] D. Kingston. *Decentralized control of multiple UAVs for perimeter and target surveillance*. PhD thesis, Brigham Young University, 2007.

[133] D. Kingston and R. Beard. UAV Splay State Configuration for Moving Targets in Wind. pages 109–128, 2007.

[134] M. Kobilarov. Cross-entropy randomized motion planning. In *RSS*, number 3, 2011.

[135] M. Krstić and H. Deng. *Stabilization of Nonlinear Uncertain Systems*. Springer-Verlag, London, 1998.

[136] V. Kumar, D. Rus, and G. S. Sukhatme. Networked Robotics. In B. Sciliano and O. Khatib, editors, *Springer Handbook of Robotics*, chapter 41, pages 943–958. 2008.

[137] H. J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, second edition, 2001. Chaps. 3–5, pp. 53–151.

[138] G. Lafferriere, a. Williams, J. Caughman, and J. Veerman. Decentralized control of vehicle formations. *Systems & Control Letters*, 54(9):899–910, Sept. 2005.

[139] E. Lalish, K. A. Morgansen, and T. Tsukamaki. Oscillatory control for constant-speed unicycle-type vehicles. In *Proc. 46th IEEE Conference on Decision and Control*, pages 5246–5251, New Orleans, LA, 2007. IEEE.

[140] A. Lambert and D. Gruyer. Safe path planning in an uncertain-configuration space. In *Proceedings of the 2003 International Conference on Robotics and Automation*, pages 85–90, 2003.

[141] L. Lapierre, D. Soetanto, and A. Pascoal. Nonsingular path following control of a unicycle in the presence of parametric modelling uncertainties. *International Journal of Robust and Nonlinear Control*, (April):485–503, 2006.

[142] S. M. Lavalle. *Planning Algorithms*. Cambridge University Press, New York, NY, 2006.

[143] J. Lawton, R. Beard, and B. Young. A decentralized approach to formation maneuvers. *IEEE Transactions on Robotics and Automation*, 19(6):933–941, Dec. 2003.

[144] M. Lemmon, T. Chantem, X. S. Hu, and M. Zyskowski. On self-triggered full-information H-infinity controllers. In *Proceedings of the 10th International Workshop on Hybrid Systems: Computation and Control*, pages 371–384, Pisa, Italy, 2007.

[145] N. Leonard and J. Graver. Model-based feedback control of autonomous underwater gliders. *IEEE Journal of Oceanic Engineering*, 26(4):633–645, 2001.

[146] X. R. Li and V. P. Jilkov. Survey of Maneuvering Target Tracking. Part I : Dynamic Models. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4): 1333–1364, 2003.

[147] X. Lin, S. L. Janak, and C. A. Floudas. A New robust optimization approach for scheduling under uncertainty: I. Bounded uncertainty. *Computers and Chemical Engineering*, 28:1069–1085, 2004.

[148] Z. Lin, B. Francis, and M. Maggiore. Necessary and sufficient graphical conditions for formation control of unicycles. *Automatic Control, IEEE Transactions on*, 50(1):121–127, 2005.

[149] K. Listmann. Consensus for formation control of nonholonomic mobile robots. In *Proceedings of the 2009 IEEE Conference on Robotics and Automation (ICRA)*, pages 3886–3891, 2009.

[150] K. Listmann, M. Masalawala, and J. Adamy. Consensus for formation control of nonholonomic mobile robots. *2009 IEEE International Conference on Robotics and Automation*, pages 3886–3891, May 2009.

[151] S.-J. Liu and M. Krstić. Stochastic source seeking for nonholonomic unicycle. *Automatica*, 46(9), Sept. 2010.

[152] S.-j. Liu, J.-f. Zhang, and Z.-p. Jiang. A Notion of Stochastic Input-to-State Stability and Its Application to Stability of Cascaded Stochastic Nonlinear Systems. *Acta Mathematicae Applicatae Sinica, English Series*, 24(1):141–156, 2008.

[153] S. Loizu, D. Dimarogonas, and K. Kyriakopoulos. Decentralized feedback stabilization of multiple nonholonomic agents. *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, pages 3012–3017 Vol.3, 2004.

[154] A. W. Long, K. C. Wolfe, M. J. Mashner, and G. S. Chirikjian. The Banana Distribution is Gaussian : A Localization Study with Exponential Coordinates. In *Proceedings of Robotics: Science and Systems*, Syndey, 2012.

[155] A. D. Luca, G. Oriolo, and C. Samson. *Feedback Control of a Nonholonomic Car-like Robot Feedback Control of a*. 1998.

[156] G. Machmani. *Minimum energy flight paths for uavs using mesoscale wind forecasts and approximate dynamic programming*. PhD thesis, Naval Postgraduate School, 2007.

[157] X. Mao. *Stochastic Differential Equations and Applications*. Horwood Publishing, Chichester, UK, 2nd edition, 2008.

[158] J.R. Marden and J.S. Shamma, Game theory and distributed control In *Handbook of Game Theory, Volume 4*, H. Peyton Young and S. Zamir (eds), Elsevier, 2013.

[159] H. Marino, M. Bonizzato, R. Bartalucci, P. Salaris, and L. Pallottino. Motion planning for two 3D-Dubins vehicles with distance constraint. *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, (1):4702–4707, Oct. 2012.

[160] S. Martinez, J. Cortes, and F. Bullo. Motion coordination with distributed information. *IEEE Control Systems*, (August):75–88, 2007.

[161] D. Q. Mayne, J. B. Rawlings, C. V. Rao, P. O. M. Scokaert, C. National, and F. Telecom. Constrained model predictive control : Stability and optimality. *Automatica*, 36(2000):789–814, 2000.

[162] M. Mazo and P. Tabuada. On event-triggered and self-triggered control over sensor/actuator networks. In *Proceedings of the 47th IEEE Conference on Decision and Control*, pages 435–440, Cancun, Mexico, 2008.

[163] M. Mazo, A. Anta, and P. Tabuada. An ISS self-triggered implementation of linear controllers. *Automatica*, 46(8):1310–1314, Aug. 2010.

[164] W. M. McEneaney. A curse-of-dimensionality-free numerical method for solution of certain HJB PDEs. *SIAM Journal on Control and Optimization*, 46(4): 1239, 2007.

[165] McGee and Hendrick. Optimal path planning with a kinematic airplane model. *AIAA Journal of Guidance, Navigation, and Control*.

[166] T. G. McGee and J. K. Hedrick. Optimal path planning with a kinematic airplane model. *Journal of Guidance, Control, and Dynamics*, 30(2):629–633, 2007. doi: 10.2514/1.25042.

[167] R. L. McNeely, R. V. Iyer, and P. R. Chandler. Tour Planning for an Unmanned

Air Vehicle Under Wind Conditions. *Journal of Guidance, Control, and Dynamics*, 30(5):1299–1306, Sept. 2007.

[168] R. L. McNeely, R. V. Iyer, and P. R. Chandler. Tour planning for an unmanned air vehicle under wind conditions. *Journal of Guidance, Control, and Dynamics*, 30(5):1299–1306, 2007. doi: 10.2514/1.26055.

[169] D. Milutinović. Utilizing Stochastic Processes for Computing Distributions of Large-Size Robot Population Optimal Centralized Control. In *Proceedings of the 10th International Symposium on Distributed Autonomous Robotic Systems*, Lausanne, Switzerland, 2010.

[170] D. Milutinović and D. P. Garg. Stochastic model-based control of multi-robot systems. Technical Report 0704, Duke University, Durham, NC, 2009.

[171] D. Milutinović and D. P. Garg. A sampling approach to modeling and control of a large-size robot population. In *Proceedings of the 2010 ASME Dynamic Systems and Control Conference*, Boston, MA, 2010. ASME.

[172] A. Molin and S. Hirche. On the Optimality of Certainty Equivalence for Event-Triggered Control Systems. *IEEE Transactions on Automatic Control*, 58(2): 470–474, 2013.

[173] R. J. Mullen, D. Monekosso, S. Barman, and P. Remagnino. Reactive Coordination and Adaptive Lattice Formation in Mobile Robotic Surveillance Swarms. *Control*.

[174] T. P. Nascimento, A. P. Moreira, and A. G. Scolari Conceição. Multi-robot nonlinear model predictive formation control: Moving target and target absence. *Robotics and Autonomous Systems*, 61(12):1502–1515, Dec. 2013.

[175] C. Navasca and A. Krener. Solution of Hamilton Jacobi Bellman equations. *Proceedings of the 39th IEEE Conference on Decision and Control (Cat. No.00CH37187)*, 1:570–574.

[176] P. Nebot and E. Cervera. Visual-aided guidance for the maintenance of multi-robot formations. *DArs*, pages 1–12.

[177] D. Necsulescu and E. Pruner. Control of nonholonomic autonomous vehicles and their formations. In *Proceedings of the 15th International Conference on Methods and Models in Automation and Robots*, pages 37–42, 2010.

[178] D. R. Nelson, D. B. Barber, T. W. McLain, and R. W. Beard. Vector field path following for miniature air vehicles. *IEEE Transanctios on Robotics and Au-*

tomation, 23(3):519–529, 2007. doi: 10.1109/TRO.2007.898976.

[179] D. R. Nelson, D. B. Barber, T. W. Mclain, S. Member, and R. W. Beard. Vector Field Path Following for Miniature Air Vehicles. 23(3):519–529, 2007.

[180] V. Nevistic and J. A. Primbs. Constrained Nonlinear Optimal Control : A Converse HJB Approach. *Control*, 1996.

[181] C. Nowzari and J. Cortés. Self-triggered coordination of robotic networks for optimal deployment. *Automatica*, 48(6):1077–1087, 2012.

[182] A. Ohsumi. Optimal search for a Markovian target. *Naval Research Logistics*, 38(4):531–554, Aug. 1991.

[183] B. Oksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, Berlin, 6th edition, 2003.

[184] R. Olfati-Saber. Flocking for Multi-Agent Dynamic Systems: Algorithms and Theory. *IEEE Transactions on Automatic Control*, 51(3):401–420, Mar. 2006.

[185] A. Palmer and D. Milutinović. A Hamiltonian Approach Using Partial Differential Equations for Open-Loop Stochastic Optimal Control. In *Proceedings of the 2011 American Control Conference*, San Francisco, CA, 2011.

[186] L. Panait and S. Luke. Cooperative Multi-Agent Learning: The State of the Art. *Autonomous Agents and Multi-Agent Systems*, 11(3):387–434, Nov. 2005.

[187] S. Panikhom and S. Sujitjorn. Numerical Approach to Construction of Lyapunov Function for Nonlinear Stability Analysis. *Res. J. App. Sci. Eng. Technol.* 4(17): 2915–2919, 2012.

[188] A. Papachristodoulou and S. Prajna. On the construction of Lyapunov functions using the sum of squares decomposition. In *Proceedings of the 41st IEEE Conference on Decision and Control.*, Las Vegas, NV, pages 3482–3487, 2002.

[189] G. Pappas and S. Simic. Consistent abstractions of affine control systems. *IEEE Transactions on Automatic Control*, 47(5):745–756, May 2002.

[190] L. E. Parker. Multiple Mobile Robot Systems. In B. Sciliano and O. Khatib, editors, *Springer Handbook of Robotics*, chapter 40, pages 921–941. Springer, 2008.

[191] V. S. Patsko and V. L. Turova. Numerical study of the "homicidal chauffeur" differential game with the reinforced pursuer. *Game Theory and Applications*,

12:123–152, 2007.

[192] T. Paul, T. R. Krogstad, J. T. Gravdahl, and E.-L. Gmbh. UAV Formation Flight using 3D Potential Field. (4):1240–1245, 2008.

[193] C. Petres, Y. Pailhas, and P. Patron. Path planning for autonomous underwater vehicles. *IEEE Transactions on Robotics*, 23(2):331-341, 2007.

[194] R. Postoyan, A. Anta, D. Nesic, and P. Tabuada. A unifying Lyapunov-based framework for the event-triggered control of nonlinear systems. In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, pages 2559 – 2564, Orlando, FL, 2011.

[195] W. B. Powell. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Wiley-Interscience, Hoboken, NJ, 2007.

[196] W. B. Powell. Perspectives of approximate dynamic programming. *Annals of Operations Research*, Feb. 2012.

[197] J. Primbs, V. Nevistić, and J. Doyle. Nonlinear optimal control: A control Lyapunov function and receding horizon perspective. *Asian Journal of Control*, 1 (1):14–24, 1999.

[198] A. Proud, M. Pachter, and J. D'Azzo. Close formation flight control. In *AIAA Guidance, Navigation, and Control Conference and Exhibit*, pages 1231–1246, Portland, OR, 1999. AIAA.

[199] C. G. Prévost, A. Desbiens, E. Gagnon, and D. Hodouin. Unmanned Aerial Vehicle Optimal Cooperative Obstacle Avoidance in a Stochastic Dynamic Environment. *Journal of Guidance, Control, and Dynamics*, 34(1):29–43, Jan. 2011.

[200] S. Quintero, F. Papi, D. J. Klein, L. Chisci, and J. P. Hespanha. Optimal UAV coordination for target tracking using dynamic programming. In *Proc. 49th IEEE Conference on Decision and Control*, 2010.

[201] S. A. P. Quintero, G. E. Collins, and P. Hespanha. Flocking with Fixed-Wing UAVs for Distributed Sensing : A Stochastic Optimal Control Approach. In *Proceedings of the 2013 American Control Conference*, pages 2028–2034, Washington, DC, 2013. IEEE.

[202] M. Rabi and J. S. Baras. Level-triggered control of a scalar linear system. In *Proceedings of the 2007 Mediterranean Conference on Control & Automation*, number 1, Athens, Greece, June 2007. IEEE.

[203] J. Reeds and L. Shepp. Optimal paths for a car that goes both forwards and backwards. *Pacific Journal of Mathematics*, 145(2):367–393, Oct. 1990.

[204] W. Ren. Trajectory tracking control for a miniature fixed-wing unmanned air vehicle. *International Journal of Systems Science*, 38(4):361–368, Jan. 2007.

[205] W. Ren and R. Beard. Trajectory Tracking for Unmanned Air Vehicles With Velocity and Heading Rate Constraints. *IEEE Transactions on Control Systems Technology*, 12(5):706–716, Sept. 2004.

[206] W. Ren and R. Beard. *Distributed consensus in multi-vehicle cooperative control: Theory and applications*. Springer Verlag, New York, NY, 2007.

[207] C. Reynolds. Flocks, Herds, and Schools: A Distributed Behavioral Model. In M. C. Stone, editor, *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pages 25–34. ACM, 1987.

[208] B. Rhoads, I. Mezi, and A. Poje. Minimum Time Feedback Control of Autonomous Underwater Vehicles. *Mechanical Engineering*, pages 5828–5834, 2010.

[209] E. Rimon and D. Koditschek. Exact robot navigation using artificial potential functions. *IEEE Transactions on Robotics and Automation*, 8(5):501–518, 1992.

[210] L. C. G. Rogers. Pathwise stochastic optimal control. pages 1–19, 2006.

[211] G. Roussos. 3D navigation and collision avoidance for nonholonomic aircraft-like vehicles. *International Journal of Adaptive Control and Signal Processing*, 24(10):900–920, 2010.

[212] G. Roussos and K. Kyriakopoulos. Decentralized and prioritized navigation and collision avoidance for multiple mobile robots. In A. Martinoli, F. Mondada, N. Correll, G. Mermoud, M. Egerstedt, M. A. Hsieh, L. E. Parker, and K. Stoy, editors, *10th International Symposium on Distributed Autonomous Robotic Systems*, pages 189–202. Springer-Verlag, 2010.

[213] G. P. Roussos, G. Chaloulos, K. J. Kyriakopoulos, and J. Lygeros. Control of multiple non-holonomic air vehicles under wind uncertainty using model predictive control and decentralized navigation functions. In *IEEE Conference on Decision and Control*, pages 1225–1230, 2008.

[214] A. Ryan, M. Zennaro, A. Howell, R. Sengupta, and J. Hedrick. An overview of emerging results in cooperative UAV control. *2004 43rd IEEE Conference on Decision and Control*, pages 602–607 Vol.1, 2004.

[215] R. Rysdyk. Unmanned aerial vehicle path following for target observation in wind. *Journal of Guidance, Control, and Dynamics*, 29(5):1092–1100, 2006.

[216] C. Samson and K. Ait-Abderrahim. Mobile Robot Control, Part 1: Feedback Control of a Nonholonomic Wheeled Cart in Cartesian Space. Technical report, Institut National de Recherche en Informatique et en Automatique, Le Chesnay, France, 1990.

[217] R. Sanfelice and E. Frazzoli. On the optimality of Dubins paths across heterogeneous terrain. *Hybrid Systems: Computation and Control*, (1):1–14, 2008.

[218] S. Särkkä. Continuous-time and continuous-discrete-time unscented Rauch-Tung-Striebel smoothers. *Signal Processing*, 90(1):225–235, Jan. 2010.

[219] S. Sarkkä. EKF/UFK Toolbox for Matlab V1.3, Available online: `http://becs.aalto.fi/en/research/bayes/ekfukf/`. 2011.

[220] R. Sattigeri, A. J. Calise, and J. H. Evers. An adaptive vision-based approach to decentralized formation control. In *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, number August, 2004.

[221] L. Sauter and P. Palmer. Path planning for fuel-optimal collision-free formation flying trajectories. In *Proceedings of the 2011 Aerospace Conference*, pages 1–10, Mar. 2011.

[222] K. Savla, E. Frazzoli, and F. Bullo. On the point-to-point and traveling salesperson problems for Dubins' vehicle. In *Proceedings of the 2005 American Control Conference*, pages 786–791, Portland, OR, 2005.

[223] K. Savla, E. Frazzoli, and F. Bullo. On the dubins traveling salesperson problems: novel approximation algorithms. In G. S. Sukhatme, S. Schaal, W. Burgard, and D. Fox, editors, *Robotics: Science and Systems II*, Philadelpha, PA, 2006. MIT Press.

[224] K. Savla, E. Frazzoli, and F. Bullo. Traveling Salesperson Problems for the Dubins Vehicle. *IEEE Transactions on Automatic Control*, 53(6):1378–1391, 2008.

[225] K. Savla, F. Bullo, and E. Frazzoli. Traveling salesperson problems for a double integrator. *IEEE Transactions on Automatic Control*, 54(4):788–793, 2009.

[226] P. Schermerhorn and M. Scheutz. Social coordination without communication in multi-agent territory exploration tasks. *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems - AAMAS '06*,

page 654, 2006.

[227] C. Schumacher. Adaptive control of UAVs in close-coupled formation flight. *Proceedings of the 2000 American Control Conference*, (June):849–853, 2000.

[228] S. K. Shah, C. D. Pahlajani, and H. G. Tanner. Probability of success in stochastic robot navigation with state feedback. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3911–3916. Ieee, Sept. 2011.

[229] S. K. Shah, C. D. Pahlajani, N. a. Lacock, and H. G. Tanner. Stochastic receding horizon control for robots with probabilistic state constraints. In *2012 IEEE International Conference on Robotics and Automation*, pages 2893–2898, Saint Paul, MN, May 2012.

[230] M. Shanmugavel, A. Tsourdos, B. a. White, and R. Ž̄bikowski. Differential Geometric Path Planning of Multiple UAVs. *Journal of Dynamic Systems, Measurement, and Control*, 129(5):620, 2007.

[231] B. Sharma and J. Vanualailai. Lyapunov Stability of a Nonholonomic Car-like Robotic System. *Nonlinear Studies*, 14(2):143 – 160, 2007.

[232] A. Singh, M. Batalin, M. Stealey, V. Chen, M. Hansen, T. Harmon, G. Sukhatme, and W. Kaiser. Mobile robot sensing for environmental applications. In *Field and service robotics*, pages 125–135. Springer, 2008.

[233] S. N. Singh, P. Chandler, C. Schumacher, S. Banda, and M. Pachter. Nonlinear Adaptive Close Formation Control of Unmanned Aerial Vehicles. *Dynamics and Control*, 10:179–194, 2000.

[234] C. Song. Numerical optimization method for HJB equations with its application to receding horizon control schemes. *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 333–338, Dec. 2009.

[235] E. D. Sontag. Input to state stability: Basic concepts and results. In P. Nistri and G. Stefani, editors, *Nonlinear and Optimal Control Theory*, chapter 3, pages 163–220. Springer-Verlag, Berlin, 2008.

[236] P. Souères and J. Boissonnat. Optimal Trajectories for Nonholonomic Mobile Robots. In *Robot motion planning and control*. Springer Berlin, pages 93–170, 1998.

[237] P. Souères, A. Balluchi, and A. Bicchi. Optimal feedback control for route tracking with a bounded-curvature vehicle. *International Journal of Control*, 74(10):

1009–1019, Jan. 2001.

[238] M. Soulignac. Feasible and Optimal Path Planning in Strong Current Fields. *IEEE Transactions on Automatic Control*, 27(1):89–98, 2011.

[239] R. F. Stengel. *Optimal Control and Estimation*. Dover Publications, New York, 1994.

[240] T. Stirling and D. Floreano. Energy-Time Efficiency in Aerial Swarm Deployment. In *Springer Tracts in Advanced Robotics: Distributed Autonomous Robotic Systems*, pages 1–12. 2013.

[241] L. D. Stone. *Theory of Optimal Search*. Academic Press, Inc., London, 1975.

[242] 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. Technical report, 1991.

[243] H. J. Sussmann. The Markov-Dubins problem with angular acceleration control. *Proc. 36th IEEE Conference on Decision and Control*, pages 2639–2643, 1997.

[244] H. J. Sussmann. The Markov-Dubins problem with angular acceleration control. In *Proceedings of 36th IEEE Conference on Decision and Control*, pages 2639–2643, San Diego, CA, Dec. 1997.

[245] R. Sutton and A. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.

[246] P. Svestka and M. H. Overmars. Probabilistic Path Planning. In J.-P. Laumond, editor, *Robot Motion Planning and Control*, chapter 5. Springer, 1998.

[247] P. Tabuada. Abstractions of Hamiltonian Control Systems. *Automatica*, 39(12): 2025–2033, 2003.

[248] P. Tabuada. Event-Triggered Real-Time Scheduling of Stabilizing Control Tasks. *IEEE Transactions on Automatic Control*, 52(9):1680–1685, 2007.

[249] P. Tabuada, G. J. Pappas, and P. Lima. Feasible Formations of Multi-Agent Systems Feasible Formations of Multi-Agent Systems '. 2001.

[250] H. Tanner and A. Kumar. Towards Decentralization of Multi-robot Navigation Functions. *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pages 4132–4137, 2005.

[251] H. Tanner and A. Kumar. *Formation stabilization of multiple agents using decentralized navigation functions*. MIT Press, 2005.

[252] H. Tanner, A. Jadbabaie, and G. Pappas. Coordination of multiple autonomous vehicles. In *IEEE Mediterranean Conference on Control and Automation*, Rhodes, Greece, 2003. IEEE.

[253] L. Techy and C. A. Woolsey. Minimum-time path-planning for unmanned aerial vehicles in steady uniform winds. *Journal of Guidance, Control, and Dynamics*, 32(6):1736–1746, 2009. doi: 10.2514/1.44580.

[254] L. Techy, C. Woolsey, and K. Morgansen. Planar path planning for flight vehicles in wind with turn rate and acceleration bounds. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 3240–3245. IEEE, 2010.

[255] B. O. Teixeira, L. a.B. Tôrres, L. a. Aguirre, and D. S. Bernstein. On unscented Kalman filtering with state interval constraints. *Journal of Process Control*, 20 (1):45–57, Jan. 2010.

[256] E. Theodorou, J. Buchli, and S. Schaal. A Generalized Path Integral Control Approach to Reinforcement Learning. *The Journal of Machine Learning Research*, 11:3137–3181, 2010.

[257] E. Theodorou and E. Todorov. Relative entropy and free energy dualities: Connections to path integral and KL control. In *Proceedings of the 2012 IEEE Conference on Decision and Control*, Maui, HI, 2012.

[258] B. Thomaschewski. Dubins' problem for the free terminal direction. *Preprint*, pages 1–14, 2001.

[259] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. The MIT Press, Cambridge, MA, 2005.

[260] U. Thygesen. A survey of lyapunov techniques for stochastic differential equations. Technical Report 18, Technical University of Denmark, 1997.

[261] E. Todorov. A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. *Proceedings of the 2005 American Control Conference*, pages 300–306.

[262] E. Todorov. Finding the Most Likely Trajectories of Optimally-Controlled Stochastic Systems. 1999.

[263] E. Todorov. General duality between optimal control and estimation. In *47th*

IEEE Conference on Decision and Control, number 5, pages 4286–4292, Cancun, Mexico, 2008. IEEE.

[264] E. Todorov. Efficient computation of optimal actions. *Proceedings of the National Academy of Sciences of the United States of America*, 106(28):11478–83, July 2009.

[265] E. Todorov. Efficient computation of optimal actions. *Proceedings of the National Academy of Sciences of the United States of America*, 106(28):11478–83, July 2009.

[266] E. Todorov. Policy gradients in linearly-solvable mdps. *Advances in Neural Information Processing Systems*, 1(1):1–10, 2010.

[267] E. Todorov and Y. Tassa. Iterative local dynamic programming. In *Adaptive Dynamic Programming and Reinforcement Learning*, 2009.

[268] S. Tonetti, M. Hehn, S. Lupashin, and R. D. Andrea. Distributed Control of Antenna Array with Formation of UAVs. In *Proceedings of the IFAC World Congress*, pages 7848–7853, 2011.

[269] J. van den Berg, P. Abbeel, and K. Goldberg. LQG-MP: Optimized path planning for robots with motion uncertainty and imperfect state information. *The International Journal of Robotics Research*, 30(7):895–913, June 2011.

[270] J. van den Berg, S. Patil, R. Alterovitz, P. Abbeel, and K. Goldberg. LQG-Based Planning, Sensing, and Control of Steerable Needles. In D. Hsu, V. Isler, J.-C. Latombe, and M. C. Lin, editors, *Springer Tracts in Advanced Robotics: Algorithmic Foundations of Robotics IX*, pages 373–389. Berlin, 2011.

[271] B. van den Broek, W. Wiegerinck, and B. Kappen. Graphical model inference in optimal control of stochastic multi-agent systems. *Journal of Artificial Intelligence Research*, 32(1):95–122, 2008.

[272] B. van den Broek, W. Wiegerinck, and B. Kappen. Optimal control in large stochastic multi-agent systems. *Adaptive Agents and Multi-Agent Systems III. Adaptation and Multi-Agent Learning*, pages 15–26, 2008.

[273] T. van den Broek. A virtual structure approach to formation control of unicycle mobile robots. In *IEEE Conference on Decision and Control*, pages 1–23, 2009.

[274] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, Amsterdam, 3rd edition, 2007.

[275] M. Velasco, P. Marti, and J. Fuertes. The self triggered task model for real-time control systems. In *Work in Progress Proceedings of the 24th IEEE Real-Time Systems Symposium*, pages 67–70, Cancun, Mexico, 2003.

[276] T. Vicsek and A. Zafeiris. Collective motion. *Physics Reports*, 517(3-4):71–140, Aug. 2012.

[277] D. D. Šiljak. *Decentralized control of complex systems*. Academic Press, Inc., San Diego, CA, 1991.

[278] A. Wächter. On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming. *Mathematical Programming*, 106(1):25–57, 2006.

[279] E. A. Wan and R. van der Merwe. Unscented Kalman Filter. In S. Haykin, editor, *Kalman Filtering and Neural Networks*, chapter 7, pages 221–282. John Wiley & Sons, Inc., New York, 2000.

[280] M. C. Wang and G. Uhlenbeck. On the theory of Brownian Motion II. *Reviews of Modern Physics*, 17(2-3):323–342, 1945.

[281] X. Wang and M. Lemmon. Event design in event-triggered feedback control systems. In *Proceedings of the 47th IEEE Conference on Decision and Control*, pages 2105–2110, Cancun, Mexico, 2008.

[282] X. Wang and M. D. Lemmon. Self-triggered Feedback Control Systems with Finite-Gain L 2 Stability. *IEEE Transactions on Automatic Control*, 54(3):452–467, 2009.

[283] X. Wang and M. D. Lemmon. Self-Triggering Under State-Independent Disturbances. *IEEE Transactions on Automatic Control*, 55(6):1494–1500, June 2010.

[284] X. Wang, V. Yadav, and S. Balakrishnan. Cooperative UAV formation flying with obstacle/collision avoidance. *Control Systems Technology, IEEE Transactions on*, 15(4):672–679, 2007.

[285] W. Wiegerinck, B. Broek, and H. Kappen. Stochastic optimal control in continuous space-time multi-agent systems. In *22nd Conference on Uncertainty in Artificial Intelligence*, Cambridge, MA, 2006.

[286] W. Wiegerinck, B. van den Broek, and B. Kappen. Optimal on-line scheduling in stochastic multiagent systems in continuous space-time. *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, page 1, 2007.

[287] T. Wik, P. Rutqvist, and C. Breitholtz. State-constrained control based on linearization of the Hamilton-Jacobi-Bellman equation. In *49th IEEE Conference on Decision and Control*, number 4, pages 5192–5197, Atlanta, GA, 2010. IEEE.

[288] R. A. Wise and R. T. Rysdyk. UAV coordination for autonomous target tracking. In *AIAA Guidance, Navigation, and Control Conference*, pages 1–22, Keystone, Colorado, 2006.

[289] W. Xi, X. Tan, and J. Baras. Gibbs sampler-based coordination of autonomous swarms. *Automatica*, 42(7):1107–1119, July 2006.

[290] M. Xue and G. S. Hornby. An Analysis of the Potential Savings from Using Formation Flight in the NAS. *AIAA Guidance, Navigation, and Control Conference*, 2012.

[291] D. Yanakiev and I. Kanellakopoulos. A Simplified Framework for String Stability Analysis in AHS. In *Proceedings of the 13th IFAC World Congress*, volume 182, pages 177–182, 1996.

[292] J. Yong. Relations among ODEs, PDEs, FSDEs, BDSEs, and FBSDEs. In *Proceedings of the 36th IEEE Conference on Decision and Control*, number December, pages 2779–2784, San Diego, CA, 1997. IEEE.

[293] C. Yoshioka and T. Namerikawa. Formation Control of Nonholonomic Multi-Vehicle Systems based on Virtual Structure. *Control*, pages 5149–5154, 2008.

[294] E. Zermelo. Über das Navigationsproblem bei ruhender oder veränderlicher Windverteilung. *Zeitschrift für Angewandte Mathematik und Mechanik (Journal of Applied Mathematics and Mechanics)*, 11(2):114–124, 1931.

[295] S. Zhu, D. Wang, and Q. Chen. Standoff tracking control of moving target in unknown wind. In *Proceedings of the 48th IEEE Conference on Decision and Control*, pages 776–781, Shanghai, China, December 15–18, 2009.

[296] S. Zhu, D. Wang, S. Member, and Q. Chen. Standoff Tracking Control of Moving Target in Unknown Wind. *Proceedings of the 2009 IEEE Conference on Decision and Control*, pages 776–781, 2009.

[297] Y. Zou, P. R. Pagilla, and R. T. Ratliff. Distributed Formation Flight Control Using Constraint Forces. *Journal of Guidance, Control, and Dynamics*, 32(1): 112–120, Jan. 2009.

Appendix A

PDF of State of Target (r, φ) after Observation used in Chapter 4

The probability density of the system state after a target observation can be found by considering the probability density function for a symmetric bivariate Gaussian in polar coordinates aligned with the direction of UAV motion θ (see Fig. 4.1(b)). Based on the last-observed position in Cartesian coordinates $[r(t^-)\cos(\varphi(t^-)),\ r(t^-)\sin(\varphi(t^-))]^T$ and the variance $\sigma^2\tau(t^-)$, this density can be found in Cartesian coordinates as

$$
\begin{aligned}
&p(\tilde{x},\tilde{y}|r(t^-),\varphi(t^-),\tau(t^-)) \\
&\quad\propto \exp\left\{-\frac{1}{2\sigma^2\tau(t^-)}\left(\left(\tilde{x}-r(t^-)\cos(\varphi(t^-))\right)^2\right.\right. \\
&\qquad\left.\left.+\left(\tilde{y}-r(t^-)\sin(\varphi(t^-))\right)^2\right)\right\} \\
&\quad= \exp\left\{-\frac{1}{2\sigma^2\tau(t^-)}\left(\tilde{x}^2+\tilde{y}^2+r(t^-)^2\right.\right. \\
&\qquad\left.\left.-2\tilde{x}r(t^-)\cos(\varphi(t^-))-2\tilde{y}r(t^-)\sin(\varphi(t^-))\right)\right\}.
\end{aligned}
$$

Reverting to the (r,φ) parameterization, with $r^2 = \tilde{x}^2 + \tilde{y}^2$, $\tilde{x}/r = \cos\varphi$, and $\tilde{y}/r = \sin\varphi$, the jump probability is

$$\begin{aligned} & p(r,\varphi | r(t^-), \varphi(t-), \tau(t-)) \\ & \quad \propto r \exp\left\{ -\frac{1}{2\sigma^2 \tau(t^-)} \left(r^2 + r(t^-)^2 \right.\right. \\ & \quad\quad \left.\left. -2 r r(t^-) \cos(\varphi - \varphi(t^-)) \right) \right\}. \end{aligned} \tag{A.1}$$

We truncate the PDF to remain within the computational domain, normalized to unity.

Appendix B

Derivations for Chapter 5

Derivation of Relative Stochastic Kinematic Model (5.9)-(5.10) for (W1)

Given a stochastic differential equation for the state $\mathbf{x} \in \mathbb{R}^n$ in the form

$$\mathrm{d}\mathbf{x}(t) = b(\mathbf{x})\mathrm{d}t + a(\mathbf{x})\mathrm{d}W(t),$$

the Itô Lemma states that the total differential of a scalar, time-independent function $f(\mathbf{x})$ is

$$\mathrm{d}\left[f(\mathbf{x})\right](t) = \left(b(\mathbf{x})\mathrm{d}t + a(\mathbf{x})\mathrm{d}W(t)\right)^{\mathsf{T}} \nabla_{\mathbf{x}} f(\mathbf{x}) + \frac{1}{2}\left(a(\mathbf{x})\mathrm{d}W(t)\right)^{\mathsf{T}} \nabla_{\mathbf{x}}^2 f(\mathbf{x}) \left(a(\mathbf{x})\mathrm{d}W(t)\right),$$

where, if $\mathrm{d}W(t)$ is of dimension k, we also have by definition that $\mathrm{d}W^{\mathsf{T}}\mathrm{d}W = I_{k\times k}\mathrm{d}t$. Applying this rule to (W1), we may obtain the total differential for $r(t) = \sqrt{(x(t))^2 + (y(t))^2}$

as

$$\begin{aligned} \mathrm{d}r(t) &= \frac{x}{r}\mathrm{d}x(t) + \frac{y}{r}\mathrm{d}y(t) + \frac{1}{2}\left(\frac{1}{r} - \frac{x^2}{r^3}\right)(\mathrm{d}x(t))^2 + \frac{1}{2}\left(\frac{1}{r} - \frac{y^2}{r^3}\right)(\mathrm{d}y(t))^2 \\ &\qquad - \frac{xy}{r^3}(\mathrm{d}x(t))(\mathrm{d}y(t)) \\ &= \left(-v\cos(\varphi) + \frac{\sigma_W^2}{2r}\right)\mathrm{d}t - \sigma_W\cos(\theta-\varphi)\mathrm{d}W_x - \sigma_W\sin(\theta-\varphi)\mathrm{d}W_y, \end{aligned} \tag{B.1}$$

where we have used the fact that $x/r = -\cos(\theta-\varphi)$ and $y/r = -\sin(\theta-\varphi)$. Similarly, since $\tan^{-1}(y/x) = \theta - \varphi + \pi$, the total differential for φ is

$$\begin{aligned} \mathrm{d}\varphi(t) &= \frac{u}{\rho_{\min}}\mathrm{d}t + \frac{y}{r^2}\mathrm{d}x(t) - \frac{x}{r^2}\mathrm{d}y(t) - \frac{xy}{r^4}(\mathrm{d}x(t))^2 + \frac{xy}{r^4}(\mathrm{d}y(t))^2 \\ &= \left(\frac{v}{r}\sin\varphi + \frac{u}{\rho_{\min}}\right)\mathrm{d}t + \frac{\sigma_W}{r}\sin(\theta-\varphi)\mathrm{d}W_x - \frac{\sigma_W}{r}\cos(\theta-\varphi)\mathrm{d}W_y. \end{aligned} \tag{B.2}$$

Since the components of the original 2D Brownian motion model are scaled with the same parameter σ_W, the noise is invariant under a rotation of the coordinate frame [105]. Defining $\mathrm{d}W_{\parallel}(t)$ and $\mathrm{d}W_{\perp}(t)$ as the increments $\mathrm{d}W_x$ and $\mathrm{d}W_y$ when viewed in a coordinate frame aligned with the direction of DV motion, we obtain (5.9)-(5.10).

Derivation of Value Iteration Equations

The following derivation of the equations for value iteration is specific to the wind model (W2). The discretization details for (W1) may be found in [10]. Denote by $\mathcal{L}^u$ the differential operator associated with the stochastic process (5.16), which, for the sake of brevity, one writes in terms of the mean drift $\mathbf{b}(\mathbf{x},u) \in \mathbb{R}^3$, the diffusion

$\mathbf{a}(\mathbf{x}) \in \mathbb{R}^{3\times 3}$ and the state vector $\mathbf{x} = [r, \varphi, \gamma]^\intercal$, as follows

$$\mathrm{d}\mathbf{x} = b(\mathbf{x}, u)\mathrm{d}t + a(\mathbf{x})\mathrm{d}W(t)$$

with the associated differential operator

$$\mathcal{L}^u = \sum_{i=1}^{3} b_i(\mathbf{x}, u)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{3} a_{ij}(\mathbf{x})\frac{\partial^2}{\partial x_i \partial x_j}.$$

The state $\mathbf{x}$ is in the domain $\mathbb{X} = \{\mathbf{x} \mid \delta \le r < r_{\max}, -\pi \le \varphi \le \pi, -\pi \le \gamma \le \pi\}$, which is semi-periodic because $[r,\ \pi, \gamma]^\intercal = [r, -\pi, \gamma]^\intercal$ and $[r,\ \varphi, -\pi]^\intercal = [r, \varphi, \pi]^\intercal$. It follows that the domain boundary is composed of two disjoint segments, i.e., $\partial\mathbb{X} = \{\mathbf{x}\ :\ r = \delta\} \cup \{\mathbf{x}\ :\ r = r_{\max}\}$.

It can be shown [137] that a sufficiently smooth $J(\mathbf{x})$ given by (5.4) satisfies

$$\mathcal{L}^u J(\mathbf{x}) + 1 = 0, \tag{B.3}$$

so that the stochastic Hamilton-Jacobi-Bellman equation for the minimum cost $V(\mathbf{x})$ over all control sequences is

$$\inf_{|u|\le 1} [\mathcal{L}^u V(\mathbf{x}) + 1] = 0. \tag{B.4}$$

This PDE has mixed boundary conditions on $\partial\mathbb{X}$. At $r = r_{\max}$, one can use reflecting boundary conditions $(\nabla V(\mathbf{x}))^\intercal \hat{n} = 0$ with the boundary normals $\hat{n}$. For the part of boundary $r = \delta$ that belongs to the target set $\mathcal{T}$, one has to use an absorbing boundary condition with $V(\mathbf{x}) = g(\mathbf{x}) \equiv 0$.

A discrete-time Markov chain $\{\xi_n, n < \infty\}$ with controlled transition probabilities from the state $\mathbf{x}$ to the state $\mathbf{y} \in \mathbb{X}$ denoted by $p(\mathbf{y} \mid \mathbf{x}, u)$ is introduced. A continuous-time approximation $\xi(t)$ to the original process $\mathbf{x}(t)$ is created by way of a state- and

control-dependent interpolation interval $\Delta t_u = \Delta t(\mathbf{x}, u) = t_{n+1} - t_n$ via $\xi(t) = \xi_n$ where $t \in [t_n, t_{n+1})$ [137]. The transition probabilities $p(\mathbf{y} \mid \mathbf{x}, u)$ then appear as coefficients in the finite-difference approximations of the operator $\mathcal{L}^u$ in (B.3). Using the so-called up-wind approximations for derivatives, the finite-difference discretizations for $J(\cdot)$ with step sizes Δr, $\Delta\varphi$, and $\Delta\gamma$ are

$$\begin{aligned} J^h(r,\varphi,\gamma) = \Delta t^u + \sum_{i=1,2} \Big\{ & p\left(r-(-1)^i\Delta r, \varphi, \gamma \mid r, \varphi, \gamma, u\right) J^h(r-(-1)^i\Delta r, \varphi, \gamma) \\ & + p\left(r, \varphi-(-1)^i\Delta\varphi, \gamma \mid r, \varphi, \gamma, u\right) J^h(r, \varphi-(-1)^i\Delta\varphi, \gamma) \\ & + p\left(r, \varphi, \gamma-(-1)^i\Delta\gamma, \mid r, \varphi, \gamma, u\right) J^h(r, \varphi, \gamma-(-1)^i\Delta\gamma) \Big\} \end{aligned} \tag{B.5}$$

where the coefficients multiplying $J^h(\cdot)$ are the respective transition probabilities, given by

$$\begin{aligned} p(r \pm \Delta r, \varphi, \gamma \mid r, \varphi, \gamma, u) &= \Delta t^u \frac{\max\left[0, (\mp v\cos(\varphi) \mp v_w \cos(\varphi+\gamma))\right]}{\Delta r}, \\ p(r, \varphi \pm \Delta\varphi, \gamma \mid r, \varphi, \gamma, u) &= \Delta t^u \frac{\max\left[0, (\pm(v/r)\sin(\varphi) \pm (v_w/r)\sin(\varphi+\gamma) \pm u/\rho_{\min})\right]}{\Delta\varphi}, \\ p(r, \varphi, \gamma \pm \Delta\gamma \mid r, \varphi, \gamma, u) &= \Delta t^u \left(\frac{\max\left[(\pm u/\rho_{\min})\right]}{\Delta\gamma} + \frac{\sigma_\theta^2}{2(\Delta\gamma)^2} \right), \end{aligned} \tag{B.6}$$

where "max" is a result of the up-wind approximation, and where Δt^u, given by

$$\begin{aligned} \Delta t^u(\mathbf{x}) = \Bigg(& \frac{|-v\cos(\varphi) - v_w\cos(\varphi+\gamma)|}{\Delta r} \\ & + \frac{|(v/r)\sin(\varphi+\gamma) + (v_w/r)\sin(\varphi+\gamma) + u/\rho_{\min}|}{\Delta\varphi} \\ & + \frac{|u/\rho_{\min}|}{\Delta\gamma} + \frac{\sigma_\theta^2}{(\Delta\gamma)^2} \Bigg)^{-1}, \end{aligned}$$

ensures that all probabilities sum to unity.

The Markov chain defined by these transition probabilities satisfies the requirement

of "local consistency," in the sense that the drift and covariance of the Markov process $\xi(t)$ are consistent with the drift and covariance of the original process, and the cost-to-go $V^h(\cdot)$ for $\xi(t)$, therefore, suitably approximates that associated with the original process. The dynamic programming equation for the Markov chain used for value iteration, is given as follows [137]:

$$V^h(\mathbf{x}) = \min_{|u|\leq 1} \left\{ \Delta t^u(\mathbf{x}, u) + \sum_{\mathbf{y}} p(\mathbf{y} \mid \mathbf{x}, u) V^h(\mathbf{y}) \right\}, \tag{B.7}$$

for all $\mathbf{x} \in \mathbb{X} \setminus \partial\mathbb{X}$. For the reflective part of the boundary, $r = r_{\max}$ (see Ref. [137, pp. 143]) is used instead of (B.7):

$$V^h(\mathbf{x}) = \sum_{\mathbf{y}} p(\mathbf{y} \mid \mathbf{x}) V^h(\mathbf{y}), \tag{B.8}$$

where $p(\mathbf{y} \mid \mathbf{x}) = 1$ for $\mathbf{y} = [r_{\max} - \Delta r, \varphi, \gamma]^\intercal$ and $\mathbf{x} = [r_{\max}, \varphi, \gamma]^\intercal$; otherwise, $p(\mathbf{y} \mid \mathbf{x}) = 0$. Finally, for those states $\mathbf{x} \in \mathcal{T}$ in the target set, it is imposed that

$$V^h(\mathbf{x}) = 0. \tag{B.9}$$

Equations (B.7)-(B.9) are used in the method of value iteration until the cost converges. From this, given the wind speed v_w, one obtains the optimal angular velocity of the DV for any relative distance r, viewing angle φ, and relative wind direction γ.