GAMES OF PURSUIT-EVASION WITH MULTIPLE AGENTS AND SUBJECT TO UNCERTAINTIES

A Dissertation
Presented to
The Academic Faculty

By

Venkata Ramana Makkapati

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Aerospace Engineering

Georgia Institute of Technology

May 2021

GAMES OF PURSUIT-EVASION WITH MULTIPLE AGENTS AND SUBJECT TO UNCERTAINTIES

Thesis committee:

Dr. Panagiotis Tsiotras, Advisor
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Yongxin Chen
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Seth Hutchinson
School of Interactive Computing
Georgia Institute of Technology

Dr. Daigo Shishika
Department of Mechanical Engineering
George Mason University

Dr. Kyriakos Vamvoudakis
School of Aerospace Engineering
Georgia Institute of Technology

Date approved: April 21, 2021

असतो मा सद्गमय। तमसो मा ज्योतिर्गमय। मृत्योर्माऽमृतं गमय॥

From the unreal lead me to the real.
From the darkness lead me to the light.
From death lead me to immortality.

— *Bṛhadāraṇyaka Upaniṣad*

कायेन वाचा मनसेन्द्रियैर्वा बुद्ध्यात्मना वा प्रकृतेः स्वभावात् ।

करोमि यद्यद् सकलं परस्मै नारायणायेति समर्पयामि ॥

Whatever I do with my body, speech, mind, sensory organs,
with my intellect, or innate natural tendencies, I offer everything to
Lord Narayana

ACKNOWLEDGMENTS

I express my most sincere gratitude to my advisor Prof. Panagiotis Tsiotras for his patience and support during my doctoral research. He reminds me of the father of the legendary Sanskrit poet Bharavi, who always wanted his son to keep aiming higher than what has already been achieved and to never rest. I cannot admire enough his ability to break down complex ideas using mathematics, and his attention to detail - qualities I hope to imbibe someday.

I am deeply thankful to Prof. Seth Hutchinson for his collaboration, and for providing pragmatic perspectives to conducting research in the areas of control and robotics. His approach to looking at a research idea using amusing yet thought-provoking examples is something that I had thoroughly enjoyed. I would like to express my gratitude to Prof. Kyriakos Vamvoudakis, Prof. Yongxin Chen, and Prof. Daigo Shishika for taking their time to serve on my thesis committee. I also thank Jack, Mehregan, Wei, Dipankar, and Kaz for their collaboration, and for being considerate labmates.

I would like to express my gratitude to my previous advisor Prof. Mangal Kothari, who has been influential in my progress from the moment I met him. I also thank Sam Miller for being an awesome flight instructor.

I am thankful to Nitansh, Ashwin, Abhijith, Vivek, Arjun, and Achyuth for being amazing roommates during my PhD, and for making my life at home less stressful than it could have been. I will forever cherish the moments we had shared. I also need to thank Ashwin specifically for introducing me to some amazing kritis in Carnatic music, hiking in national parks, and Futurama.

I am indebted to Tanmay Rajpurohit for taking the role of a *Guru* in my life. His discourses on topics ranging from control systems, mathematics, quantum mechanics to history, philosophy, and spirituality are invaluable. I sincerely thank him for trying to show me what this is all about, and I don't know if I have realized it yet!

I thank Hemanth, Sasanka, Nagu babai, Lakshmi pinni, Ramana babai, Padma pinni, Murthy babai, Sita pinni, Vasudevan Comandur and his family (Sandhya aunty and Vandhu) for their unending support. I would like to express my gratitude to Vinodhini. She has been a true friend since the time we met in IIT Kanpur.

I will be forever indebted to my family (Amma, Naanna, and Yasaswi) for their support in every way possible; not just during the PhD journey, but from the moment I started to exist. There is no way to thank them for their struggles to get me to this point of academic excellence, and I humbly offer my pranams. I also offer my pranams to my grandparents Shri Makkapati Srinivasa Sastry (late), Shrimati Makkapati Venkata Lakshmi (late), Shri Vishnubhatla Lakshmi Narayana (late), and Vishnubhatla Venkata Lakshmi for their continued blessings in this world and from beyond.

Finally, I would like to reminisce this *Subhashita* (wise words) that discusses the process of how one acquires knowledge.

आचार्यात् पादमादत्ते पादं शिष्यः स्वमेधया ।
पादं सब्रह्मचारिभ्यः पादं कालक्रमेण च ॥

One fourth from the teacher,
one fourth from his own intelligence,
one fourth from classmates, and
one fourth only with time.

While I expressed gratitude to my teachers and classmates, I can only but bow down to *Lord Mahakaleshwara*, who is time itself.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Over the past decade, there have been constant efforts to induct unmanned aerial vehicles (UAVs) into military engagements, disaster management, weather monitoring, and package delivery, among various other applications. With UAVs starting to come out of controlled environments into real-world scenarios, uncertainties that can be either exogenous or endogenous play an important role in the planning and decision-making aspects of deploying UAVs. At the same time, while the demand for UAVs is steadily increasing, major governments are working on their regulations. There is an urgency to design surveillance and security systems that can efficiently regulate the traffic and usage of these UAVs, especially in secured airspaces. With this motivation, the thesis primarily focuses on airspace security, providing solutions for safe planning under uncertainties while addressing aspects concerning target acquisition and collision avoidance.

In this thesis, we first present our work on solutions developed for airspace security that employ multiple agents to capture multiple targets in an efficient manner. Since multi-pursuer multi-evader problems are known to be intractable, heuristics based on the geometry of the game are employed to obtain task-allocation algorithms that are computationally efficient. This is achieved by first analyzing pursuit-evasion problems involving two pursuers and one evader. Using the insights obtained from this analysis, a dynamic allocation algorithm for the pursuers, which is independent of the evader's strategy, is proposed. The algorithm is further extended to solve multi-pursuer multi-evader problems for any number of pursuers and evaders, assuming both sets of agents to be heterogeneous in terms of speed capabilities.

Next, we consider stochastic disturbances, analyzing pursuit-evasion problems under stochastic flow fields using forward reachability analysis, and covariance steering. The problem of steering a Gaussian in adversarial scenarios is first analyzed under the framework of general constrained games. The resulting covariance steering problem is solved

numerically using iterative techniques. The proposed approach is applied to the missile endgame guidance problem. Subsequently, using the theory of covariance steering, an approach to solve pursuit-evasion problems under external stochastic flow fields is discussed. Assuming a linear feedback control strategy, a chance-constrained covariance game is constructed around the nominal solution of the players. The proposed approach is tested on realistic linear and nonlinear flow fields. Numerical simulations suggest that the pursuer can effectively steer the game towards capture.

Finally, the uncertainties are assumed to be parametric in nature. To this end, we first formalize optimal control under parametric uncertainties while introducing sensitivity functions and costates based techniques to address robustness under parametric variations. Utilizing the sensitivity functions, we address the problem of safe path planning in environments containing dynamic obstacles with an uncertain motion model. The sensitivity function based-approach is then extended to address game-theoretic formulations that resemble a "fog of war" situation.

CHAPTER 1

INTRODUCTION

1.1 Motivation

Coordination strategies for unmanned aerial vehicles (UAVs) has been an active area of research especially in the realm of multi-agent systems over the past decade [1, 2, 3, 4, 5], having numerous applications, including agriculture, aerial surveying, fire detection, disaster management, weather monitoring, and commercial product delivery (eg: Amazon Prime Air in Figure 1.1(a)). UAVs already play a major role in military engagement scenarios, and their use as part of swarm tactics (encirclement, coordinated attack, search and rescue, perimeter defense) promises to change future battlefield operations. However, the deployment of small UAVs for various missions is currently facing endurance limitations. To this end, some of the solution techniques that are currently under investigation include extending their range capabilities and reducing their fuel consumption by leveraging environmental winds, landing on and riding along with ground vehicles. The latter technique involves rendezvous with a moving target that can improve the functionality of these UAVs.

(a) © Amazon.com, Inc.

(b) © 2018 Government of Singapore

Figure 1.1: (a) Prime Air is a future delivery system from Amazon designed to safely get packages to customers in 30 minutes or less using UAVs (b) A drone-catcher pursues drones that enter secured airspaces and captures them using a net. Counter-drone technologies involving multiple UAVs are among the solutions for smart airbases for future.

More recently, an analysis on the current commercial UAV market suggests that their use is expected to grow manyfold over the coming years, as aerial drones are becoming a household product, used for recreational and industrial purposes alike[1]. These advancements suggest an urgent need to explore designs for airspace safety systems that can regulate the traffic and usage of UAVs in a large scale (eg: Drone Catcher[2] in Figure 1.1(b)). Similarly, we are slowly but surely inching towards the world of fully autonomous self-driving vehicles, and before these vehicles become a reality, they will have to interact with human-driven vehicles in traffic, as well as pedestrians. Since human drivers can be extremely unpredictable, such autonomous vehicles have to plan for worst case scenarios to guarantee safety.

In all the above mentioned problems, which involve multi-agent systems, coordination strategies obtained by formulating them as a multi-player pursuit-evasion (PE) game offer solutions that address collision avoidance, surveillance and target acquisition [6, 7, 8, 9]. However, obtaining solutions to such multi-player differential games still remains a challenge. Additionally, all of the above instances require the autonomous agents to operate in an environment that is uncertain and adversarial. In the case of UAVs, environmental disturbances (such as winds) can greatly affect the system's performance, and in turn decide the fate of a mission.

The major goal of this research is to develop novel solutions for controlling multiple unmanned vehicles having a common goal under the influence of external disturbances and uncertainties. The external disturbances can include winds (in the case of aerial vehicles), water currents (in the case of marine vehicles), or traffic flow (in the case of self-driving cars). The difference in the actual and perceived behavior of other adversarial agents (asymmetric information) in the environment account to the uncertainties part of the research. More specifically, the thesis presents decentralized strategies for the players in pursuit-evasion games involving multiple agents and uncertainties.

[1] http://www.businessinsider.com/commercial-uav-market-analysis-2017-8

[2] https://www.straitstimes.com/singapore/saf-announces-plans-to-tap-tech-consolidate-bmt

1.2 Literature Survey

The seminal work of Rufus Isaacs on differential games paved the way for the study of pursuit-evasion (PE) problems [10]. In [10], some techniques to study multi-player PE games were briefly discussed. Following Isaacs, various formulations of the PE game were analyzed. An extensive amount of literature is now available, and a recent survey on zero-sum PE games with multiple agents is available in Ref. [11].

The two-pursuer/one-evader problem has been previously discussed by H. Kelley in [12], where some aspects of two-on-one team tactics were identified. Bhattacharya and Hutchinson analyzed the problem subject to visibility constraints, and provided approximate schemes to construct the set of initial pursuer configurations from which capture is guaranteed [13]. The differential game involving two pursuers and one evader with linear dynamics and quadratic cost was investigated for Nash equilibrium solutions in Ref. [14]. Partitions of the state-space were identified and categorized, which are similar to the degenerate and the non-degenerate regions discussed later on in this paper. A linear differential game formulation restricting the motion of the players to a straight line with the two pursuers coordinating to reduce the miss distance was extensively studied by several researchers [15, 16]. An algorithm to numerically construct level sets of the value function with fixed final time was also discussed in Ref. [17]. A case of two identical inertial pursuers (second order dynamics) pursuing a non-inertial evader (first order dynamics) was studied by Levchenkov and Pashkov [18]. Hagedorn and Breakwell considered the problem of a faster evader that must pass between two pursuers [19]. Finally, a version of the relay pursuit problem discussed in this paper was previously presented by Sun and Tsiotras, along with a suboptimal strategy [20].

Evasion from a group of pursuers is a subset of the class of multi-player PE games. Classical results include those of Pshenichnyi, who provided a sufficient condition for successful evasion from a group of homogeneous pursuers [21], Blagodatskikh [22] and

Chernous'ko [23], among many others. Specifically, Chernous'ko showed that an evader can avoid point capture from any number of pursuers having a lower speed. The PE differential game involving many pursuers and one evader has also been investigated subject to fixed terminal time, integral constraints and different pay-off models [24, 25].

Obtaining closed-form optimal strategies for the players in general multi-agent PE games using Hamilton-Jacobi-Isaacs (HJI) equation formulations is elusive, owing to the curse of dimensionality. Jang and Tomlin proposed some control strategies obtained from direct differentiation of a given value function, but these strategies are sub-optimal [26]. An extension to this problem, which assumes that the evader is more agile than the pursuers, was studied by Zak [27]. Oyler et al. [28] studied planar PE games in the presence of obstacles by constructing dominant regions for each player. Some limitations of capturing a faster evader were proposed, and heuristic group pursuit strategies were presented in [29, 30, 31, 32, 33]. In another version of the group pursuit problem, a group of faster, yet less agile, pursuers against a slower, but more agile evader, was solved by Bopardikar et al. [34]. Group pursuit problems involving general dynamics were studied in [35, 36, 37]. A probabilistic variant of group pursuit problems was investigated in [38] using a greedy policy. Lastly, relay group pursuit using dynamic Voronoi diagrams was studied in [39, 40].

The literature for multi-pursuer multi-evader (MPME) problems, where there are players from two teams encountering each other with conflicting motives [41], is rather limited compared to its multi-pursuer single-evader (MPSE) problems. In most cases, some form of heuristic is introduced in order to make the problem tractable. Ge et al. [42] proposed three approaches, which include hierarchical decomposition, moving horizon hierarchical decomposition, and cooperative control. Li et al. [43] also explored a hierarchical approach, while Jin and Qu [44] proposed a heuristic task allocation algorithm. Extensions to the MPME problem includes problems with incomplete information [45], nonlinear dynamics [46], and a mix of continuous and discrete observations [47]. However, finding scalable algorithms which can be implemented in real-life MPME scenarios is still an open

problem [48, 49, 50, 51].

Most of the aforementioned solution approaches can be categorized into formal methods that provide optimal (or sub-optimal) strategies for the players. The learning paradigm propelled alternate approaches to obtain model-free solutions for some of the aforementioned pursuit-evasion formulations using deep neural networks [52, 53]. In the recent past, techniques based on Q-learning [54, 55], policy gradients [56, 57], and actor-critic methods [58] were explored to analyze multi-agent pursuit-evasion problems. In [56] and [59], formulations with closed domains were proposed to solve the game with a superior evader using deep reinforcement learning. A survey of the reinforcement learning based techniques for multi-agent systems can be found in Ref. [60], that discusses the associated challenges. Some of the challenges include defining a "good" learning goal, lack of theoretical and convergence guarantees, scalability, and the exploration-exploitation trade-off.

Planning under environmental disturbances, such as wind fields and uncertain currents, is a necessity for technological solutions employing many aerial and underwater autonomous vehicles. Traditionally, such disturbances are assumed to be stochastic, and there is a vast amount of work available for planning under stochastic uncertainties, both endogenous and exogenous. To this end, stochastic pursuit-evasion games have received a great deal of attention by many researchers over the years, who have proposed various formulations and numerical techniques [61, 62, 63, 64].

One of the most promising approaches for solving deterministic pursuit-evasion (PE) games involves reachability and level set based analysis [65, 66]. These have been applied in aerospace applications such as for the construction of safety envelope [67]. Sun et al. derived capture conditions and open-loop strategies for agents in multi-player PE problems with dynamic flow fields using a reachability based approach [66]. The work by Sun et al. employs forward reachable sets by solving the level set equations. Under the assumption that the evader is slower than the pursuer in the one-pursuer-one-evader game, the approach led to a simplified capture condition, stating that the optimal capture time is the minimum

time taken by the pursuer's reachability set to contain the evader's reachability set. In the stochastic realm, forward reachability based analysis is a relatively new idea with limited previous work, and the system dynamics was mostly assumed to be linear [68, 69, 70, 71].

Prior work on stochastic dynamic games considered continuous-time linear systems with players having noisy measurements. In one of the early works, Speyer discussed the outcomes of a linear stochastic differential game in a defense setting with a missile and a radar [72]. The radar's objective is to minimize the uncertainty in the current state of the missile using a filter, and the missile strives to maximize the same, while also trying to arrive at a target point with certain accuracy. Further, the missile has linear dynamics and adds additive noise to its control in order to confuse its adversary. Subsequently, zero-sum LQ differential game theory with noise-corrupted measurements was applied to study the pursuit-evasion setting involving a homing missile and an evading aircraft [73]. Quasi-linearization was used to solve the nonlinear two-point boundary-value problem with closed-form solutions. At the same time, Castañón and Athans derived feedback strategies for a two-person Stackelberg game with quadratic performance cost and linear dynamics [74]. Kumar and van Schuppen considered different observations and costs for the two players [75]. They obtained strategies for the players assuming that one of the players had a "spy" creating an instance of asymmetric information. Bley and Stear studied discrete stochastic dynamic games in $\mathbb{R}$ [76]. Yavin analyzed various pursuit-evasion problems in the stochastic setting and provided sufficient conditions for the players' optimal strategies [77, 78]. Bernhard and Colomb provided a solution for the rabbit and the hunter game in a partial information setting, using dynamic programming [79].

In recent years, stochastic games with different information structures, where the state transition and observation equations are linear, have been extensively studied [80, 81, 82, 83, 84, 85]. Rajpurohit et al. studied a two-player stochastic differential game with a nonlinear dynamical model over an infinite time horizon [86].

Owing to the fact that a Gaussian distribution can be fully defined using its first two mo-

ments, a class of stochastic optimization problems were decomposed into their respective mean and covariance steering problems [87, 88]. The mean steering problem is essentially a deterministic dynamic game. The necessary and sufficient conditions for the existence of a solution to the discrete-time LQ dynamic game was provided by Pachter and Pham, along with a closed-form solution [89].

The problem of steering the state of a stochastic dynamic system from a given initial Gaussian distribution to a desired one is referred to as the covariance steering problem. The idea of covariance control has its genesis in the 1980s [90]. The problem of finite-horizon covariance control in continuous time was however analyzed only recently by Chen et al. [91, 92, 93]. In Ref. [88], state chance constraints were introduced to the covariance control problem in the context of path planning with static obstacles and system uncertainties. The approach was subsequently modified to deal with general nonlinear dynamics [94], and was applied to spacecraft control [95, 96, 97]. Previous works have shown that these solutions can be seen in the context of LQG with a particular set of weights which can be solved in terms of LMIs [98, 99, 100].

Optimal trajectory planning and feedback control are two critical and interrelated components required to achieve autonomous flight. Traditionally, a nominal trajectory is computed preflight by solving an optimization problem constrained by the vehicle's dynamical model, containing estimated atmospheric and aerodynamic parameters. The guidance law is then given as the sum of this nominal input with a feedback term, which is added so that the trajectory is robust to parametric uncertainties and random disturbances. It follows that, in a clear division of labor, the open-loop nominal control defines the nominal trajectory and the feedback control reduces sensitivity to uncertainty. Take for example the Apollo direct entry guidance, which was designed as part of the Apollo program and later adapted for the Mars Science Laboratory (MSL) and Mars 2020 entry guidance [101, 102, 103]. A nominal profile of bank angle commands are first set to determine the nominal entry path, and then feedback gains are computed as a function of the adjoint system integrated

backwards along the nominal trajectory.

In this thesis, we propose, instead, to reduce trajectory dispersion due to parametric uncertainties by selection of the nominal trajectory. The nominal control then serves a dual role of optimizing the given performance metric and reducing dispersion. Early work on trajectory sensitivity design include those of Winsor and Roy [104], who developed a technique to design controllers that provide assurance for system performance under mathematical modeling inaccuracy. The feasibility of the technique was established with appropriate simulation results. Following that work, several approaches including sensitivity-reduction for linear regulators, using increased-order augmented system [105], modification of weighting matrix [106], feedback [107, 108], and an augmented cost function [109, 110], were all thoroughly analyzed. The approach of using an augmented cost function was further tested on the linear quadratic regulator (LQR) problem, which was later applied for active suspension control in passenger cars [110].

With the work on trajectory sensitivity design mostly restricted to analyzing linear systems, more recent approaches under the title desensitized optimal control (DOC) considered sensitivities to address dispersion under modeling uncertainties in general nonlinear optimal control problems. The work by Seywald et al. makes use of sensitivity analysis to obtain an optimal open-loop trajectory that is insensitive to first-order parametric variations for general optimal control problems [111, 112]. The proposed approach elevates the parameters of interest to system states, and defines a binary sensitivity function that provides the first-order variation in the states at time t, given the variation in the states at some time t' ($t' \leq t$). An appropriate sensitivity cost is added to the existing cost function, and the dynamics of the binary sensitivity function is augmented in the system dynamics to solve the resulting optimal control problem. The approach was later extended to optimal control problems with control constraints [113], and it was used to solve the Mars pinpoint landing problem [114]. Some extensions to the landing problem include considering uncertainties in atmospheric density and aerodynamic characteristics [115], and the

use of direct collocation and nonlinear programming [116]. The process of constructing desensitized trajectories can be also understood as a way to impose smoothness through the regularization of the sensitivities [117].

Probabilistic methods are a popular choice to address planning in uncertain dynamic environments [118]. In the past, Gaussian processes have been employed to model uncertainty and to obtain safe trajectories [119, 120, 121, 122]. Techniques involving POMDPs [123, 124], occupancy grids [125], intent-based threat estimation [126], replanning [127, 128], and feedback coupled with estimation [129] are many variants in the class of probabilistic methods. Alternatively, reachability analysis [130], artificial potential fields (APFs) [131], and barrier functions [132] have also been utilized. Our approach fundamentally differs from the above techniques in its formulation; while prior work widely employed probabilistic techniques, the proposed formulation in this paper is deterministic. In order to address planning under limited sensing and feedback capabilities, we provide safe open-loop trajectories that have guarantees on optimality by treating the uncertainty to be parametric in nature, and by examining sensitivities with respect to parameter variations.

Lastly, we address planning in adversarial scenarios with asymmetric information using sensitivity based techniques. This is a game-theoretic extension to the aforementioned safe optimal planning under parametric uncertainties. Games with asymmetric/incomplete information have been the subject of interest with probabilistic methods being the most popular choice for analysis [133]. In Bayesian games, for instance, each player assumes probability distributions over the random variable denoting the uncertain information and the beliefs of the other players to arrive at an equilibrium solution that maximizes its expected rewards [134]. It has been shown that lack of common knowledge would result in a non-zero-sum game even when the original reward structure of the players corresponds to a zero-sum game [135]. In order to characterize Nash equilibria in a game with asymmetric information, Nayyar et al. constructed an equivalent symmetric game using the common information among the players [136]. Subsequently, an algorithm to compute such Nash

equilibria by solving a sequence of linear equations was proposed in [137]. The approach was later applied to address security of cyber-physical systems under resource constraints on the players [138]. An algorithm to obtain fixed points for the infinite-horizon version of the game was proposed by Vasal et al. [139].

In the realm of pursuit-evasion problems, the notion of detectability was considered to be an outcome of lack of information [140, 141]. Antoniades et al. used heuristics to investigate multi-agent games in which a team of pursuers detect and capture a team of evaders [45]. Estimation-based approaches to address pursuit-evasion with information uncertainty were first put forth by Mizukami and Tews [142]. The approach was extensively analyzed to obtain solutions for various types of orbital pursuit-evasion games [143, 144, 145, 146]. Furthermore, linear pursuit-evasion games with bounded controls and information delay was analyzed by Shinar and Glizer [147]. A control estimation technique to solve linear pursuit-evasion with uncertain relative dynamics is presented by Cavalieri et al. [148]. A particular formulation where the evader observes just the initial conditions while the pursuer observes the instantaneous relative distance with additive noise was analyzed by Hexner et al. [149]. A discrete-time version of the pursuit-evasion game with incomplete information was analyzed by Gurel-Gurevich [150]. Finally, Huang et al. provided some epistemic models to examine pursuit-evasion games with incomplete information [151].

1.3 Contributions and Thesis Outline

In this thesis, pursuit-evasion scenarios involving multiple pursuers and multiple evaders are first analyzed using geometric methods, and in these problems, the players' dynamics are assumed to be deterministic. Next, we analyze a class of linear quadratic stochastic dynamic games using the theory of general constrained games and covariance steering. Subsequently, we analyze pursuit-evasion under uncertainties by considering single-pursuer single-evader problems in stochastic flow fields. The proposed solution approach utilizes forward reachable sets and covariance control approaches. Finally, the uncertainties are

assumed to be parametric in nature and to this end, we initially develop sensitivity-based solutions for optimal control problems, which are then extended to differential games with asymmetric information. Thus, the work in this thesis can be broadly organized into three sections, and the contributions in each section are discussed below.

1.3.1 Geometric Methods for Multi-Pursuer Multi-Evader Problems

Solving exactly a multi-player dynamic game necessitates the solution of a complicated partial differential equation, whose dimensionality increases with the number of players. The problem is simplified by following a dynamic "divide and conquer" approach, where at every time instant each evader is assigned to a set of pursuers based on the instantaneous positions of all the players. This means that instead of solving the full multi-agent pursuit-evasion game involving n pursuers and m evaders, we solve a series of m many/against-one pursuit-evasion games. This leads to decentralized, although likely suboptimal, solutions. In this work we assume that the pursuers are faster compared to the evaders and they follow simple navigation laws (pure pursuit or constant bearing).

In chapter 2, the simplest case of multi-pursuer single-evader (MPSE) problem that involves two pursuers and one evader is discussed. The basic idea of the proposed methodology is to determine which of the pursuers is the most critical one and which pursuer does not affect the outcome of the game and thus may be eliminated from the analysis of the problem. This leads to the notions of degenerate and non-degenerate regions for each PE game. The determination of these regions is a key aspect towards the solution of multi-player games, since they delineate the regions in which one pursuer acting alone can capture the evader without the help of its team-mates, from the regions in which some form of cooperation/coordination between the pursuers is necessary for optimal capture. Also, we extend our work to cases where one of the two pursuers is stationary, motivated by relay-pursuit scenarios [40, 20]. From an application point of view, these strategies result in a trade-off between time and resources. For instance, employing more than one pursuer

may reduce capture time, but the deployment of more pursuers require additional resources (e.g., fuel, communication bandwidth, etc). The main contributions of chapter 2 are listed below.

1. The regions of non-degeneracy are identified in two-pursuer one-evader problems for cases where the pursuers follow either a constant bearing strategy, or a pure pursuit strategy, or a relay pursuit strategy.

2. The optimal evading strategies for the case of constant bearing and pure pursuit are identified and it is established, for the first time, that in both cases, when the problem is non-degenerate, the solution involves simultaneous capture.

3. A competitive sub-optimal strategy is suggested for pure pursuit and a comparative study is provided for the case of identical pursuers to demonstrate this claim.

4. An optimal evading strategy is derived for relay-pursuit scenarios, along with the corresponding switching condition.

In chapter 3, the concept of regions of non-degeneracy is generalized to MPSE problems using the concept of active/redundant pursuers. In the case of constant bearing, and assuming that the evader can follow any strategy, a dynamic task allocation algorithm is proposed for the pursuers. The algorithm is based on the well-known Apollonius circle and allows the pursuers to allocate their resources in an intelligent manner while guaranteeing the capture of the evader in minimum time. For the case of pure pursuit, the algorithm is modified using the counterpart of the Apollonius circle leading to an Apollonius closed curve. The main contributions of chapter 3 are listed below.

1. In the MPSE settings, the characteristics of the optimal evading strategies in both cases (constant bearing and pure pursuit) are derived, and it is established that the optimal evading strategy depends only on the initial conditions of those pursuers that finally capture (simultaneously) the evader.

2. A framework to characterize the pursuers into active and redundant in MPSE settings using Apollonius circles (for constant bearing) and Apollonius curves (for pure pursuit) is provided.

3. An algorithm to identify the status of a pursuer, given the instantaneous positions of all the players in an MPSE setting is presented. This algorithm allows us to perform a dynamic task allocation of the pursuers that ensures the evader's capture in minimum time, under any evading strategy.

4. Assuming both teams (pursuers and evaders) are homogeneous, a task allocation algorithm is proposed for the pursuers in multi-pursuer multi-evader (MPME) settings. The algorithm is scalable for any number of pursuers and evaders.

5. Finally, the task allocation algorithm for MPME problems with pursuers following a constant bearing is extended to scenarios with heterogeneous teams.

1.3.2 Covariance Control and Reachability Based Approaches for Pursuit-Evasion Games with Stochastic Disturbances

Stochastic games, introduced by Shapley in 1953, deal with instances where a stochastic process is jointly controlled by two players, a *controller* and a *stopper*, along with an underlying payoff function that is common to both players [152]. The stopper tries to maximize the payoff function, while the controller strives to minimize it. In chapter 4, we address a class of linear-quadratic (LQ) stochastic dynamic games in discrete-time with finite-time horizon. It is assumed that the players have perfect measurement of the state at each time instant and that the initial state is sampled from a given Gaussian distribution. First, the problem of steering the covariance in an LQ game setting without any constraints is analyzed, and the associated saddle point equilibrium is identified. Subsequently, the problem of steering the initial distribution to a specified terminal distribution (which is also Gaussian) under adversarial situations, which can be categorized as a general constrained

game (GCG) [153], is considered. This constrained covariance steering game is relevant in the context of stochastic pursuit-evasion and has applications in spacecraft rendezvous [154], collision avoidance[155], and understanding behavioral patterns in nature [156].

It is important to note that the attitude of a player towards its opponent's constraint influences the outcome of the GCG [153]. Instances where the players are indifferent to couple constraints of their opponents can be found in mobile networks and finance [153, 157]. For example, in the situation where there are multiple mobile carriers competing to maximize the received power in a series of time slots, the networks are also subjected to a minimum expected throughput. Note that in a given time slot, only one carrier is successful, and the overall success of a network depends on the actions of all other mobiles, indicating coupled constraints with players being indifferent to their opponents' constraints. In the case where a player's main goal is to prevent the opponent from meeting its constraint, his attitude is to be understood as being aggressive. Analyzing the scenario where the stopper has an aggressive attitude towards the controller's constraint is beyond the scope of this thesis. The main contributions of chapter 4 are listed below.

1. A novel LQ formulation for driving a Gaussian to a given terminal distribution under an adversarial setting is introduced. The adversary is assumed to be indifferent to the controller's terminal constraint which is unique to the literature on covariance steering.

2. It is shown that the proposed game theoretic formulation can be decomposed into two independent games, mean steering and covariance steering games, which makes the problem tractable.

3. The existence of equilibrium solutions is discussed for both unconstrained and constrained versions of the games.

4. A condition in terms of relative controllability is identified in the mean steering game with controller constraints for discrete systems. A simple Jacobi procedure for find-

ing saddle points is introduced to solve the one-sided constrained covariance steering game, assuming a linear feedback control structure.

5. The missile endgame guidance scenario is revisited, while assuming a process noise in the system, to demonstrate the proposed approach for one-sided constrained covariance steering game.

In chapter 5, we consider two-agent PE problems with both agents traversing a stochastic flow field. It is assumed that both agents have speed constraints, and the pursuer is superior to the evader in terms of its speed capabilities. Initially, a forward reachability analysis is performed while considering only the drift term in the flow field to obtain the nominal trajectories for the agents. Assuming a linear feedback control architecture, we then formulate a discrete-time chance-constrained covariance game about the players' nominal trajectories, which is solved using the standard Gauss-Seidel method, to obtain closed-loop controls for both players. The main contributions of chapter 5 are listed below.

1. A chance-constrained covariance steering game is formulated to solve single-pursuer single-evader games in general non-linear stochastic flow fields.

2. Assuming a linear feedback control law, Gauss-Seidel iteration technique is employed to obtain the closed-loop gains of the players.

3. The approach is demonstrated on realistic flow fields, including an approximate wind field model for Hurricane Hugo [158].

1.3.3 Sensitivity Regularization Techniques for Planning under Parametric Uncertainties

The tension between optimality and safety is often evident in robotics—particularly for applications that have stringent performance requirements—under conditions for which uncertainties in sensing, environment models, and control effectiveness are unavoidable [159, 160, 161, 162]. For all but the simplest applications, optimal solutions tend to bring

the robot dangerously close to the operational safety margins. For example, it is well known that the shortest path for a mobile robot in a polygonal environment lies in the visibility graph which implies that the optimal path would contact the obstacles while traversing the path [163]. While in practice it is typical to perturb paths slightly such that they do not reach the constraint boundaries, this safety strategy raises a number of significant questions: How should one perform these perturbations? How should one balance the cost of violating constraints against reduced performance? And, perhaps most importantly, how can one provide a principled evaluation of the effects of uncertainty with respect to the trade-offs between optimality and safety, and adjust the path to optimally balance between the two? It is this latter question that we address in chapter 6.

In chapter 6, we consider uncertainties to be parametric in nature, where the nominal value of the uncertain parameter is available. Using sensitivity functions [164], we first capture the variations in the constraint function under parametric variations, defined as *constraint sensitivity*. The variations are then weighted using a relevance function to obtain the *relevant constraint sensitivity* (RCS), and construct a regularizer that captures the risk of constraint violation. The characteristics of the regularizer are discussed by analyzing its performance in simple path planning problems. We then evaluate the proposed technique on path planning problems in environments containing up to ten dynamic obstacles having uncertain velocities.

Furthermore, an approach to desensitize the cost for an optimal control problem with fixed final time is presented in chapter 6. To this end, we recall that the co-states in an optimal control problem are a measure of the sensitivity of the value function with respect to the states along the optimal trajectory [165, 166]. It is proven that the co-states indeed capture sensitivity of the *cost-to-go* function with respect to perturbations in the state given *any* prescribed control law $u(t)$, not just the optimal one. The main contributions of chapter 6 are listed below.

1. A sensitivity function based regularizer is proposed to minimize dispersion in the

trajectory or state at a particular time under open-loop control with parametric uncertainties.

2. Relevance functions and employed to efficiently construct a sensitivity function based risk measure, which when penalized provides "safe" trajectories with a lower chance for constraint violations under parametric uncertainties.

3. It is proven that for any given control law, the co-states of an optimal control problem capture the first-order variations of the cost-to-go function given the variations in the system states.

4. An approach to desensitize (reduce the variations of) the cost with respect to parametric variations is presented. Also, it is shown that the co-states and the sensitivity matrices (in Ref. [112]) are indeed related.

In chapter 7, we consider two-player zero-sum differential games with orthogonal constraints, which are specific to a player and do not depend on the control inputs of the player's adversaries [153]. In the proposed game with asymmetric information, it is assumed that the minimizing player is at an information disadvantage, which corresponds to the lack of exact information about a set of model parameters. It is assumed that the minimizing player knows the nominal values of those parameters. The maximizing player is assumed to be completely aware of the environment, representing a form of "home-field advantage", manifested as having the exact values of the model (environment) parameters. In this paper, sensitivity functions are used to construct a regularizer that captures the risk of constraint violation for the minimizing player. The proposed approach is demonstrated on two-player pursuit-evasion games with an uncertain dynamic obstacle. The pursuer is assumed to be the player with incomplete information about the motion of the obstacle in the environment, while the evader is fully aware of the obstacle's motion model. The main contributions of chapter 7 are listed below.

1. The theory of desensitized optimal control is extended to address differential games with asymmetric information.

2. A sensitivity function based approach is presented to generate conservative trajectories for the player which has an information disadvantage in two-player differential games with asymmetric information.

3. The proposed approach is demonstrated on pursuit-evasion games with an uncertain dynamic obstacle.

Finally, chapter 8 presents some future directions for the research presented in this thesis along with concluding remarks.

CHAPTER 2
PURSUIT-EVASION SCENARIOS INVOLVING TWO PURSUERS AND ONE EVADER

2.1 Background

Consider a group of n agents (pursuers) guarding a given area of interest. The objective of the agents is to pursue and intercept m (where typically $m \leq n$) intruders (or evaders) that may be detected in this area. Some of the relevant questions that arise while solving this problem include:

1. Which pursuer(s) should go after which evader(s)?
2. How many pursuers should chase each intruder (evader) in order to capture it in the shortest time possible?
3. What is the shortest time-to-capture, given the fact that the evaders are intelligent and will try to postpone capture indefinitely?

Obtaining the answers to the previous questions in their most general form is elusive at this point. Solving exactly a multi-player dynamic game such as the one considered above will necessitate the solution of a high-dimensional partial differential equation, whose dimensionality increases with the number of players ($n+m$). In order to proceed and mitigate this problem, the following assumptions are made in this work.

A1: The pursuers are faster compared to the evaders.

A2: The pursuers follow simple navigation laws (pure pursuit or constant bearing strategies).

The rationale behind these assumptions is as follows. Under assumption A1, a pure pursuit or a constant bearing strategy guarantees capture. Also, these two pursuit strategies highlight the available information to the pursuer in order to capture the evader. A constant bearing (CB) strategy is known to be efficient when the pursuer knows the instantaneous position and velocity of the evader [167]. On the other hand, an individual pursuer that has access only to the evader's instantaneous position can, at best, employ a pure pursuit (PP) strategy [167]. Furthermore, both of these strategies are easy to execute, and they have been implemented successfully in various aerial defense systems. A more detailed discussion on navigation laws for different information structures can be found in Ref. [168].

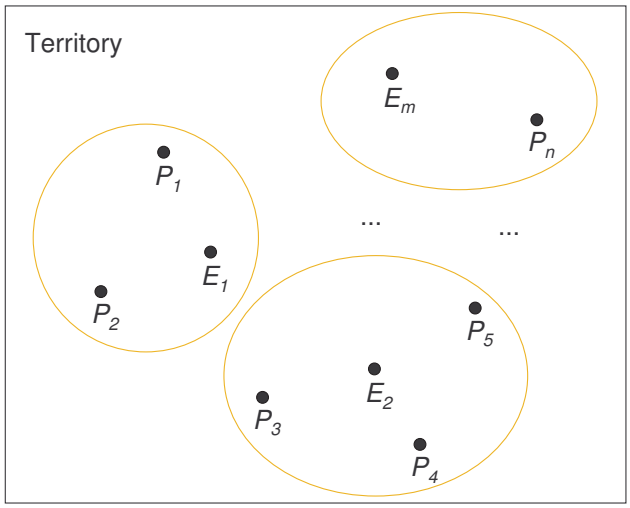


Figure 2.1: Divide and conquer in multi-pursuer multi-evader problems: P_i denotes the i^{th} pursuer, and E_j denotes the j^{th} evader

A potential approach to solve complicated multi-pursuer multi-evader (MPME) problems is a *dynamic* "divide and conquer" approach (see Figure 2.1), where the pursuers are divided into several groups corresponding to the evader they pursue at each instant of time. In essence, such divide and conquer strategies formulate the original MPME problem as a sequence of several (simpler) multi-pursuer single-evader (MPSE) problems [169, 42]. This approach leads to decentralized, although likely suboptimal, solutions. By analyzing the associated MPSE problems for the cases of CB and PP, one may arrive at an efficient

dynamic task-allocation algorithm of pursuers to evaders. The simplest case of problems involving MPSE scenarios is the case of two pursuers against one evader, which is the focus of this chapter.

2.2 Problem Formulation

Consider a pursuit-evasion (PE) problem with two pursuers and one evader in the plane. The objective of the pursuers is that at least one of them enters the evader's capture zone, assumed here to be a disk of radius $\epsilon > 0$ centered at the current position of the evader, while the objective of the evader is to avoid or delay capture as long as possible. The subscripts 1 and 2 will be used for the two pursuers (P_1 and P_2), while the subscript E will be used for the evader. The equations of motion for all the players involved in the game are given below.

$$\dot{x}_1 = u_1 \cos\theta_1, \quad \dot{y}_1 = u_1 \sin\theta_1, \tag{2.1a}$$

$$\dot{x}_2 = u_2 \cos\theta_2, \quad \dot{y}_2 = u_2 \sin\theta_2, \tag{2.1b}$$

$$\dot{x}_E = v \cos\theta_E, \quad \dot{y}_E = v \sin\theta_E, \tag{2.1c}$$

where $p_1 = (x_1, y_1)$, $p_2 = (x_2, y_2)$, and $p_E = (x_E, y_E)$ denote the positions of pursuer P_1, pursuer P_2, and the evader E, respectively. Similarly, $\theta_1,\ \theta_2,\ \theta_E \in (-\pi, \pi]$ denote the control inputs of the players, u_1, u_2 and v are the speeds (constant) of P_1, P_2, and E, respectively, with $\min\{u_1, u_2\} > v$. The game evolves in the six-dimensional state-space, $[x_1, y_1, x_2, y_2, x_E, y_E]^\top \in \mathbb{R}^6$.

Problem: Find the optimal control input for the evader, $\theta_E \in (-\pi, \pi]$ that maximizes the time of capture t_c in the following cases:

CB: The two pursuers follow a constant bearing strategy;

PP: The two pursuers follow a pure pursuit strategy;

R-CB: Two identical ($u_1 = u_2$) pursuers follow a relay pursuit strategy with the active pursuer employing constant bearing;

R-PP: Two identical ($u_1 = u_2$) pursuers follow a relay pursuit strategy with the active pursuer employing pure pursuit.

In all these cases, note that the control inputs of the pursuers, θ_1 and θ_2, depend solely on the instantaneous states of the players.

2.3 The Regions of Non-degeneracy

Assuming each of the pursuers follows either a constant bearing strategy or a pure pursuit strategy, the two-pursuer one-evader problem may result in a degenerate case. A degenerate case is one in which only one of the pursuers is sufficient to capture the evader in minimum time. In degenerate problems, the presence of one of the two pursuers is inconsequential and the problem can be treated as a one-against-one PE problem. For instance, if one of the pursuers (say P_2) is very far away from the evader (E), then P_2 does not play a role in the solution of the problem and the evader's optimal strategy is a pure evasion from P_1. Similarly, if P_2 is very close to E then P_2 dominates, and the optimal evading strategy in this case will be a pure evasion from P_2. However, there exists a region of initial positions for P_2 for which a form of coordination with P_1 ensues, and an optimal evading strategy (other than pure evasion from P_1 or P_2) needs to take into consideration the presence of both pursuers.

Since we are interested in studying the effects of adding a second pursuer (P_2) to the problem, it follows from the preceding discussion that given the initial positions of the first pursuer (P_1) and the evader (E) along with the speed capabilities of all the three players, it is important to find the set of initial positions of P_2 for which: 1) the optimal evading strategy is a pure evasion from P_1, and P_2 plays no role in the solution of the problem ($\mathcal{D}_2$ - Degenerate region with respect to P_2); 2) the optimal evading strategy if the evader follows

pure evasion from P_2, and P_1 plays no role in the solution ($\mathcal{D}_1$ - Degenerate region with respect to P_1); 3) the optimal evading strategy is not a pure evasion from any of the two pursuers, and both P_1 and P_2 play a role in the problem ($\mathcal{N}$ - Non-degenerate region). We can compute these regions for the cases CB and PP as follows.

2.3.1 Constant Bearing Strategy

Without of loss of generality, assume that the initial positions of P_1 and E are such that $p_1(0) = (0, 0)$ and $p_E(0) = (d, 0)$, respectively ($d \neq 0$). The capture time, assuming that the evader follows a pure evasion strategy from P_1, is given by

$$t_f = \frac{d}{u_1 - v}. \tag{2.2}$$

The point of capture is $C = (u_1 t_f, 0)$. Define now the circle $\mathcal{C}_1 = \{\mathsf{x} \in \mathbb{R}^2 : \|\mathsf{x} - C\| = u_2 t_f\}$. The circle $\mathcal{C}_1$ is an isochrone that contains the set of initial positions for P_2 that guarantees capture exactly at time t_f (at location C) under a constant bearing strategy for the given initial position of the evader and its heading, assuming that the evader is non-maneuvering. The center of this circle is the capture point C and its radius is $u_2 t_f$. If the initial position of P_2 (following a constant bearing strategy) lies inside $\mathcal{C}_1$, then P_2 can capture the evader in a time less than t_f for the given initial position of the evader and its heading. If the initial position of P_2 lies outside $\mathcal{C}_1$, then it cannot capture the evader in a time less than or equal to t_f, which means that the presence of P_2 is inconsequential to the solution of the problem, and the optimal evading strategy is pure evasion from P_1. As a result, in the case of CB, the circle $\mathcal{C}_1$ and its exterior constitutes the degenerate region with respect to P_2, and $\mathcal{C}_1$ is the boundary between $\mathcal{D}_2$ and $\mathcal{N}$.

Next, the degenerate region with respect to P_1 can be obtained by looking at the locus of the initial points of P_2 such that a pure evasion from P_2 would result in simultaneous capture of the evader by both P_1 and P_2. In this regard, consider the Apollonius circle

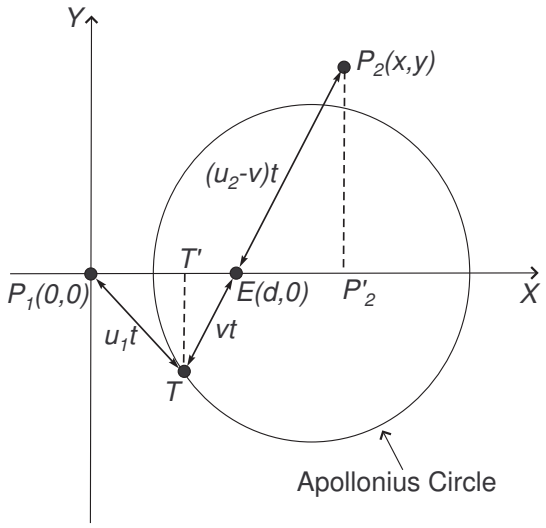


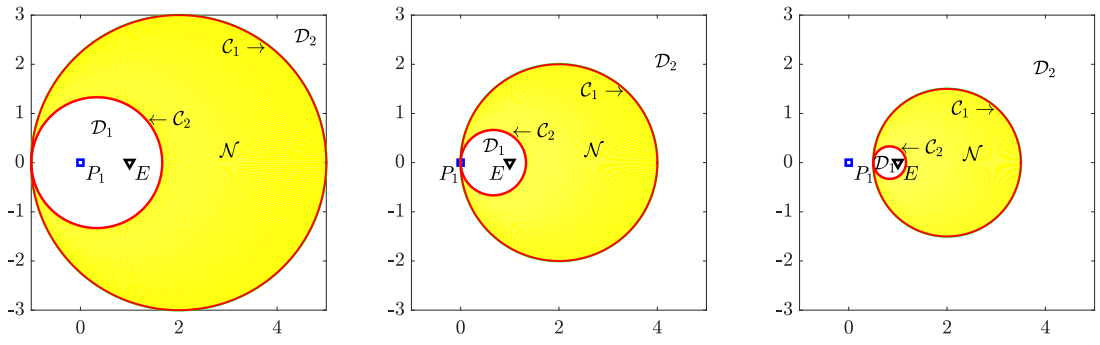
Figure 2.2: A limiting case scenario of the degeneracy with respect to P_1 for the case of CB

($\mathcal{A}$) corresponding to P_1 and E, whose center is at $(x_a, y_a) = (u_1^2 d/(u_1^2 - v^2), 0)$, and its radius is $r_a = u_1 v d/(u_1^2 - v^2)$ [10]. The circle constitutes capture points for P_1 for a given constant heading of the evader. Clearly, finding the set of initial points of P_2 such that the evader hits a point on the Apollonius circle under pure evasion from P_2 provides the locus of interest; see Figure 2.2. This can be achieved in the following manner. Consider a point T on $\mathcal{A}$ in its parametric form $T = (x_a + r_a \cos\phi, r_a \sin\phi)$ (note that $y_a = 0$) and let (x, y) be the initial position of P_2 such that it hits the evader at T. From the geometry of the problem, as defined in Figure 2.2, and since the triangles $\triangle ETT'$ and $\triangle EP_2P_2'$ are similar, it follows that

$$\frac{x - d}{d - x_a - r_a \cos\phi} = \frac{y}{r_a \sin\phi} = \frac{u_2 - v}{v}. \tag{2.3}$$

The coordinates x, y can then be given as

$$x = \left(d + \frac{(u_2 - v)(d - x_a)}{v}\right) - r_a \left(\frac{u_2 - v}{v}\right) \cos\phi, \qquad y = r_a \left(\frac{u_2 - v}{v}\right) \sin\phi. \tag{2.4}$$



(a) $u_2 = 1.5$ (b) $u_2 = 1$ (c) $u_2 = 0.5$

Figure 2.3: Regions of degeneracy and non-degeneracy for the case of CB: $u_1 = 1$, $v = 0.5$, and $\epsilon = 0$ (point capture).

Under this parametric representation, it can be realized that the set of points (x, y) form a circle (call it $\mathcal{C}_2$) with its center at $(d + (u_2 - v)(d - x_a)/v, 0)$ and radius $r_a(u_2 - v)/v$. It is understood that if P_2 lies inside $\mathcal{C}_2$, then the evader will get captured by P_2 under pure evasion, before it hits the Apollonius circle $\mathcal{A}$, i.e., P_1 does not play a role in the solution of the problem. Hence, in the CB case the circle $\mathcal{C}_2$ and its interior constitutes the degenerate region with respect to P_1, and $\mathcal{C}_2$ acts as the boundary between $\mathcal{D}_1$ and $\mathcal{N}$. Finally, $\mathbb{R}^2 \backslash (\mathcal{D}_1 \cup \mathcal{D}_2)$ constitutes the region of non-degeneracy.

The geometry of these regions can be visualized in Figure 2.3 which shows the regions of degeneracy and non-degeneracy for three different cases, where $u_2 = 1.5, 1, 0.5$. As can be seen in this figure, the evader's initial position is at $(1, 0)$ with P_1 located at $(0, 0)$. The speeds of E and P_1 are $v = 0.5$ and $u_1 = 1$, respectively. In Figure 2.3(a), P_2 is faster compared to P_1, and P_1 is inside circle $\mathcal{C}_2$. The converse is observed in Figure 2.3(c) when P_2 is slower than P_1. It can be observed that the region of non-degeneracy increases with the speed of P_2, which suggests that given the speeds of P_1 and E, adding a second pursuer which has higher speed would enable cooperation among the pursuers in a larger region, and vice versa.

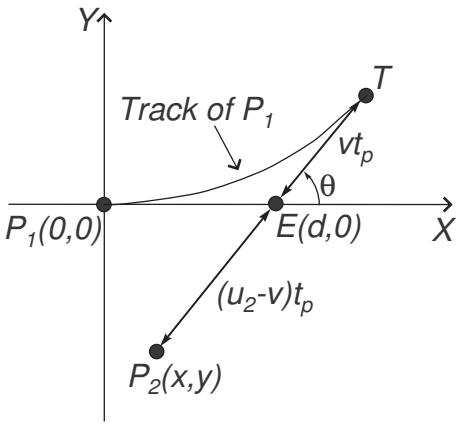
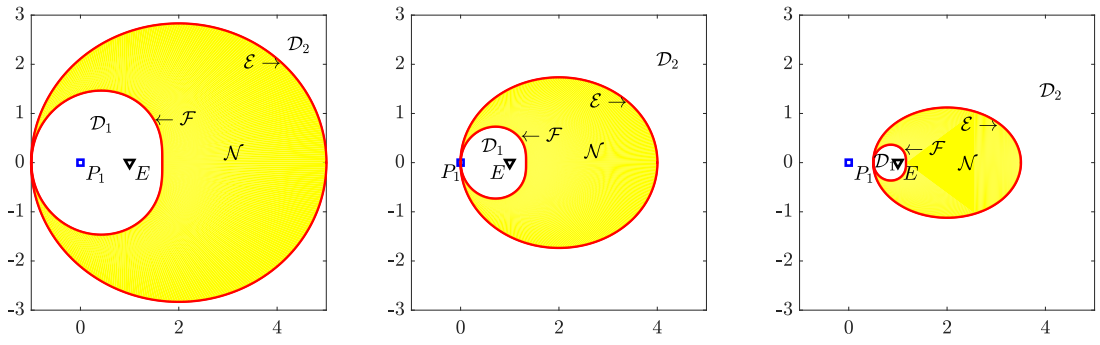


Figure 2.4: A limiting case scenario of the degeneracy with respect to P_1 for the case of PP.

2.3.2 Pure Pursuit Strategy

To compute the non-degenerate region in this case, first define the ellipse, $\mathcal{E} = \{\mathsf{x} \in \mathbb{R}^2 : \|\mathsf{x} - p_E(0)\| + \|\mathsf{x} - p'_E\| = 2u_2 t_f\}$, where $p'_E = \left(\left[\frac{u_1 + v}{u_1 - v}\right] d, 0\right)$. The ellipse $\mathcal{E}$ is an isochrone that contains the set of initial positions for a pursuer that guarantees capture at time t_f (at location C) under a pure pursuit strategy for a given initial position of the evader and its heading, assuming that the evader is non-maneuvering. In the literature, $\mathcal{C}_1$ and $\mathcal{E}$ are called t_f-isochrones [167]. It can be seen that $\mathcal{E}$ is an ellipse centered at C having the initial position of the evader at one of its foci. For any initial position of a pursuer (following a pure pursuit strategy) inside $\mathcal{E}$, the capture time is less than t_f. If the initial position of P_2 (following a constant bearing strategy) lies inside $\mathcal{C}_1$, then it can capture the evader in a time less than t_f. Therefore, if the initial position of P_2 lies outside $\mathcal{E}$, then P_2 cannot capture the evader in a time less than or equal to t_f i.e., P_2's presence has no strategic significance and the optimal evading strategy is a pure evasion from P_1. Consequently, in the case of PP, the ellipse $\mathcal{E}$ and its exterior constitutes the degenerate region with respect to P_2, and $\mathcal{E}$ is the boundary between $\mathcal{D}_2$ and $\mathcal{N}$.



(a) $u_2 = 1.5$ (b) $u_2 = 1$ (c) $u_2 = 0.5$

Figure 2.5: Regions of degeneracy and non-degeneracy for the case of PP: $u_1 = 1$, $v = 0.5$, and $\epsilon = 0$ (point capture).

The degenerate region with respect to P_1 in this case can be obtained from the relation

$$t_p = \frac{r_o(u + v\cos\theta)}{u_1^2 - v^2}, \qquad u_1 \neq v, \tag{2.5}$$

that provides the capture time for a pursuer that follows a pure pursuit strategy, assuming that the evader is non-maneuvering (constant heading) [167]. Here r_o is the initial distance between the pursuer and the evader (= d for P_1 and E), and θ is the evader's heading measured with respect to the line-of-sight from the pursuer to the evader. Using this relation, and following an approach similar to the one in subsection 2.3.1, the locus of the initial points of P_2 such that a pure evasion from P_2 would result in simultaneous capture of the evader by both P_1 and P_2 can be obtained. Consider now the case of an evader following a pure evasion strategy from P_2 with the heading θ, as shown in Figure 2.4.

Let the initial position of P_2 be (x, y). Assuming that the evader gets captured by both P_1 and P_2 at T (see Figure 2.4), it follows from Equation 2.5 that

$$x = d - (u_2 - v)d\cos\theta\left(\frac{u_1 + v\cos\theta}{u_1^2 - v^2}\right), \qquad y = -(u_2 - v)d\sin\theta\left(\frac{u_1 + v\cos\theta}{u_1^2 - v^2}\right). \tag{2.6}$$

Under this parametric representation, the locus of interest can be obtained, which is a closed

curve ($\mathcal{F}$) around the initial position of the evader. If P_2 lies inside $\mathcal{F}$, then the evader will get captured by P_2 under pure evasion, before E is captured by P_1. And hence, in the case of PP, the closed curve $\mathcal{F}$ and its interior constitutes the degenerate region with respect to P_1, and $\mathcal{F}$ acts as the boundary between $\mathcal{D}_1$ and $\mathcal{N}$. Finally, in the case of PP, the set $\mathbb{R}^2 \backslash (\mathcal{E} \cup \mathcal{F})$ constitutes the region of non-degeneracy.

The regions of degeneracy and non-degeneracy observed in this case are depicted using the previous example in subsection 2.3.1. Note that the ellipse $\mathcal{E}$ is contained in the circle $\mathcal{C}_1$, and on the other hand, the closed curve $\mathcal{F}$ contains $\mathcal{C}_2$. This is a consequence of the fact that the pursuers use pure pursuit, as they lack information about the evader's speed. As a result, the regions of non-degeneracy are smaller compared to their counterparts in subsection 2.3.1. This observation further supports the fact that the information structure plays a crucial role in problems involving cooperation among agents.

2.4 Optimal Evading Strategies Against Two Pursuers

2.4.1 Constant Bearing Strategy

As per the formulation in section 2.2, it can be seen that the game evolves in the six-dimensional state-space. However, the problem formulation can be reduced to the two-dimensional state-space in the following manner. Consider the relative distances between the evader and each of the pursuers ($r_1 - p_1$, $r_2 - p_2$), and the corresponding LoS angles (φ_1, φ_2), are shown in Figure 2.6. Using Equation 2.1, the dynamics of the relative distances and the LoS angles can be expressed as

$$\dot{r}_1 = v\cos(\theta_E - \varphi_1) - u_1\cos(\theta_1 - \varphi_1), \quad \dot{\varphi}_1 = \frac{1}{r_1}\left[v\sin(\theta_E - \varphi_1) - u_1\sin(\theta_1 - \varphi_1)\right], \tag{2.7a}$$

$$\dot{r}_2 = v\cos(\theta_E - \varphi_2) - u_2\cos(\theta_2 - \varphi_2), \quad \dot{\varphi}_2 = \frac{1}{r_2}\left[v\sin(\theta_E - \varphi_2) - u_2\sin(\theta_2 - \varphi_2)\right]. \tag{2.7b}$$

Furthermore, it is assumed that the pursuers follow a constant bearing strategy and hence the LoS for a given pursuer does not rotate, i.e., $\dot{\varphi}_1 = 0$ and $\dot{\varphi}_2 = 0$. That is, $\varphi_1(t) = \varphi_{10}$, $\varphi_2(t) = \varphi_{20}$, for all $t \geq 0$, where φ_{10} and φ_{20} are the LoS angles at the initial time $t = 0$. Therefore, r_1 and r_2 are the only states that have to be taken into consideration to solve for the optimal evading strategy.

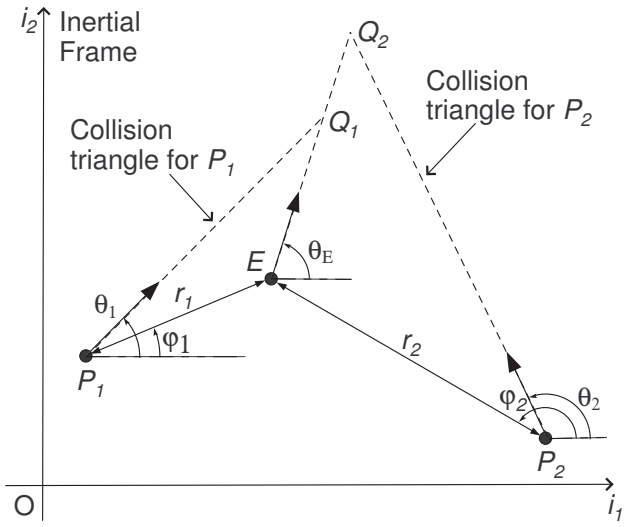


Figure 2.6: Schematics of the proposed pursuit-evasion problem with pursuers following a constant bearing strategy

We are dealing with a time maximization problem subject to the dynamics

$$\dot{r}_1 = v\cos(\theta_{_E} - \varphi_{10}) - u_1\cos(\theta_1 - \varphi_{10}), \tag{2.8a}$$

$$\dot{r}_2 = v\cos(\theta_{_E} - \varphi_{20}) - u_2\cos(\theta_2 - \varphi_{20}). \tag{2.8b}$$

Note that θ_1, θ_2 are functions of $\theta_{_E}$. Given $\theta_{_E}$, they can be determined at each instant of time, using Equation 2.7, and the relation $\dot{\varphi}_1 = \dot{\varphi}_2 = 0$. Therefore,

$$v\sin(\theta_{_E} - \varphi_{10}) = u_1\sin(\theta_1 - \varphi_{10}), \qquad v\sin(\theta_{_E} - \varphi_{20}) = u_2\sin(\theta_2 - \varphi_{20}). \tag{2.9}$$

Each of the above relations have two possible solutions for θ_1 or θ_2, given $\theta_{_E}$, and each

pursuer chooses the solution for which $\dot{r}_1 < 0$ or $\dot{r}_2 < 0$, respectively. The initial conditions are $r_1(0) = r_{10} = \|p_E(0) - p_1(0)\|$ and $r_2(0) = r_{20} = \|p_E(0) - p_2(0)\|$. The terminal condition for capture is

$$\Psi(r_1(t_c), r_2(t_c)) = \min\{r_1(t_c), r_2(t_c)\} - \epsilon = 0. \tag{2.10}$$

The Hamiltonian for this problem can be expressed as

$$\begin{aligned} H(r_1, r_2, \lambda_1, \lambda_2, \theta_E) = -1 &+ \lambda_1 \left[v \cos(\theta_E - \varphi_{10}) - u_1 \cos(\theta_1 - \varphi_{10})\right] \\ &+ \lambda_2 \left[v \cos(\theta_E - \varphi_{20}) - u_2 \cos(\theta_2 - \varphi_{20})\right], \end{aligned} \tag{2.11}$$

where λ_1 and λ_2 are the co-states. The corresponding adjoint equations are given by

$$\dot{\lambda}_1 = 0, \qquad \dot{\lambda}_2 = 0, \tag{2.12}$$

and therefore $\lambda_1(t) = c_1$, $\lambda_2(t) = c_2$, for $t \in [0, t_c]$, where c_1 and c_2 are constants. The transversality conditions are given by

$$\lambda_1(t_c) = \nu \frac{\partial \Psi}{\partial r_1}\Big|_{t=t_c}, \quad \lambda_2(t_c) = \nu \frac{\partial \Psi}{\partial r_2}\Big|_{t=t_c}, \quad H(t_c) = 0, \tag{2.13}$$

where $\nu \in \mathbb{R}$. Since the Hamiltonian has no explicit dependency on time, it follows that $H(t) = 0$ for all $t \in [0, t_c]$. Note that the terminal condition is not fully differentiable and it can be written as

$$\frac{\partial \Psi}{\partial r_1}\Big|_{t=t_c} = \frac{r_2 - \min\{r_1, r_2\}}{r_2 - r_1}, \qquad \frac{\partial \Psi}{\partial r_2}\Big|_{t=t_c} = \frac{r_1 - \min\{r_1, r_2\}}{r_1 - r_2}. \tag{2.14}$$

At $r_1(t_c) = r_2(t_c) = \epsilon$, the partial derivatives are undefined. Using Pontryagin's minimum

principle, the following expression is obtained,

$$\lambda_1 \left[-v \sin(\theta_E - \varphi_{10}) + u_1 \sin(\theta_1 - \varphi_{10}) \frac{\partial \theta_1}{\partial \theta_E} \right] + \lambda_2 \left[-v \sin(\theta_E - \varphi_{20}) + u_2 \sin(\theta_2 - \varphi_{20}) \frac{\partial \theta_2}{\partial \theta_E} \right] = 0, \tag{2.15}$$

where from Equation 2.9,

$$\frac{\partial \theta_1}{\partial \theta_E} = \frac{v \cos(\theta_E - \varphi_{10})}{u_1 \cos(\theta_1 - \varphi_{10})}, \qquad \cos(\theta_1 - \varphi_{10}) \neq 0, \tag{2.16a}$$

$$\frac{\partial \theta_2}{\partial \theta_E} = \frac{v \cos(\theta_E - \varphi_{20})}{u_2 \cos(\theta_2 - \varphi_{20})}, \qquad \cos(\theta_2 - \varphi_{20}) \neq 0. \tag{2.16b}$$

Since λ_1, λ_2 are constants, θ_1, θ_2 and their partials from Equation 2.16 are dependent only on θ_E, we can conclude from Equation 2.15 that the optimal heading of the evader θ_E is constant in time, and hence the headings of the pursuers are constant as well.

Theorem 1. *For the CB time-optimal control problem, given the initial positions of P_1 and E, if P_2 initially lies in the non-degenerate region ($\mathcal{N}$), then the optimal control strategy of the evader involves simultaneous capture by P_1 and P_2.*

Proof. First, consider the case when $r_1(t_c) = \epsilon < r_2(t_c)$, i.e, only P_1 captures the evader at the final time. In this case, from Equation 2.13 and Equation 2.14, it follows that $\lambda_2(t_c) = 0$. Subsequently, Equation 2.12 implies that $c_2 = 0$. Then from Equation 2.11 and Equation 2.15, we have

$$-1 + c_1 \left[v \cos(\theta_E - \varphi_{10}) - u_1 \cos(\theta_1 - \varphi_{10}) \right] = 0, \tag{2.17}$$

$$c_1 \left[-v \sin(\theta_E - \varphi_{10}) + u_1 \sin(\theta_1 - \varphi_{10}) \frac{\partial \theta_1}{\partial \theta_E} \right] = 0, \tag{2.18}$$

If $c_1 \neq 0$, it leads to a contradiction in Equation 2.17. Therefore,

$$-v \sin(\theta_E - \varphi_{10}) + u_1 \sin(\theta_1 - \varphi_{10}) \frac{\partial \theta_1}{\partial \theta_E} = 0. \tag{2.19}$$

From Equation 2.16a, $\sin(\theta_1 - \varphi_{10})\cos(\theta_E - \varphi_{10}) - \sin(\theta_E - \varphi_{10})\cos(\theta_1 - \varphi_{10}) = 0$, which implies that $\sin(\theta_E - \theta_1) = 0$. Further analysis leads to $\theta_E^* = \varphi_{10}$, i.e, the optimal strategy is a pure evasion from P_1. This is the solution for a degenerate case of the problem. It has been proven that in the non-degenerate case, with this evasion strategy, P_2 will capture the evader prior to P_1, leading to a contradiction. Similarly, the strategy when $r_2(t_c) = \epsilon < r_1(t_c)$ turns out to be a pure evasion from P_2. But since P_1 is closer to the evader, it lies inside the circle of equal time-to-capture corresponding to P_2, and therefore P_1 reaches the evader before P_2, again leading to a contradiction. Hence, the optimal evading strategy in the non-degenerate case should involve $r_1(t_c) = r_2(t_c) = \epsilon$ namely, simultaneous capture. □

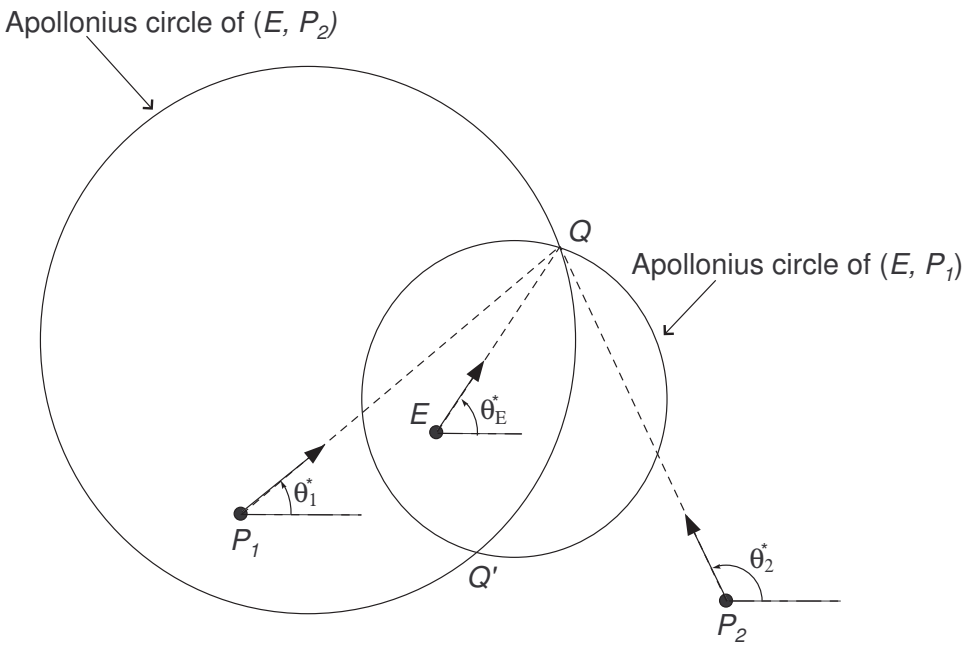


Figure 2.7: A schematic of finding the optimal heading of the evader using Apollonius circles.

Since the optimal heading θ_E^* is constant, and involves simultaneous capture in the non-degenerate case, it is easy to obtain the heading using the well-known Apollonius circles [10]; see Figure 2.7. The Apollonius circles of the pairs (E, P_1) and (E, P_2) at the initial time can be constructed from the players' initial positions. If the problem is non-degenerate, then there always exist two intersection points, Q and Q' as shown in

Figure 2.7. During optimal play, the evader should head towards one of the intersection points, namely, the one that is farther away. If both the points are equidistant, then the evader can choose either point. This completes the analysis on the optimal evading strategy for CB.

2.4.2 Pure Pursuit Strategy

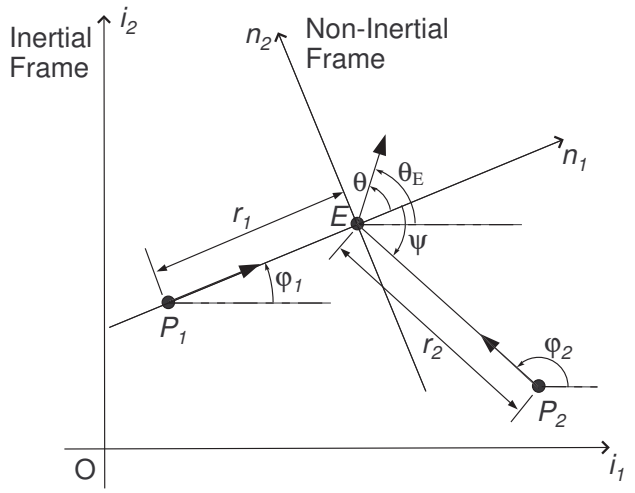


Figure 2.8: Schematics of the proposed pursuit-evasion problem with pursuers following a pure pursuit strategy

In this case, the problem can be examined in the three-dimensional state-space that makes the analysis simpler. A schematic of the geometry of the proposed pursuit-evasion problem is shown in Figure 2.8. First, we translate the problem into a rotating/non-inertial frame with the origin fixed on the evader (E) and with the x-axis along the line joining P_1 and E. The velocity vector of P_1 is along the x-axis, as it follows a pure pursuit strategy. In this frame P_1 is restricted to move only along the x-axis. The positions of the players expressed in polar coordinates are given by $p_1 = (r_1, \pi)$, $p_2 = (r_2, \psi)$, and $p_E = (0, 0)$, $-\pi < \psi \leq \pi$. Since the pursuers follow a pure pursuit strategy, their headings are along their corresponding LoS directions, i.e., $\theta_1 = \varphi_1$, $\theta_2 = \varphi_2$; see Figure 2.8. The angle between the velocity vectors of P_1 and E is $\theta = \theta_E - \varphi_1$. The rotation rate of the non-

inertial frame is given by

$$\dot{\varphi}_1 = \frac{v \sin\theta}{\|p_E - p_1\|} = \frac{v \sin\theta}{r_1}. \tag{2.20}$$

In the reduced state-space, the number of states is only three, and the corresponding equations of motion are given by

$$\dot{r}_1 = -u_1 + v\cos\theta, \tag{2.21a}$$

$$\dot{r}_2 = -u_2 - v\cos(\psi - \theta), \tag{2.21b}$$

$$\dot{\psi} = \frac{v}{r_2}\sin(\psi - \theta) - \frac{v}{r_1}\sin\theta. \tag{2.21c}$$

The initial conditions for the states are $r_1(0) = \|p_E(0) - p_1(0)\|, r_2(0) = \|p_E(0) - p_2(0)\|, \psi(0) = \pi - \varphi_{10} + \varphi_{20}$, where φ_{10} and φ_{20} are now the initial headings of P_1 and P_2, respectively, which can be obtained from the initial positions of the players. The terminal condition remains the same as in Equation 2.10.

The problem statement is then to find the optimal control $\theta^*(t)$ that maximizes the capture time t_c given the equations of motion in Equation 2.21, and the given initial conditions and terminal conditions. It is assumed that the initial conditions are such that the problem is non-degenerate for the given speeds of the players. Otherwise, the pursuit strategy for the evader is pure evasion from either pursuer.

The Hamiltonian for this problem can be written as

$$\begin{aligned} &H(r_1, r_2, \psi, \lambda_1, \lambda_2, \lambda_3, \theta) = \\ &-1 + \lambda_1(-u_1 + v\cos\theta) + \lambda_2[-u_2 - v\cos(\psi - \theta)] + \lambda_3\left[\frac{v}{r_2}\sin(\psi - \theta) - \frac{v}{r_1}\sin\theta\right], \end{aligned} \tag{2.22}$$

where λ_1, λ_2, and λ_3 are the co-states, and satisfy the adjoint equations (dropped for brevity). Since $\psi(t_c)$ is not specified and is free, the transversality conditions are given

by

$$\lambda_1(t_c) = \nu \frac{\partial \Psi}{\partial r_1}\Big|_{t=t_c}, \quad \lambda_2(t_c) = \nu \frac{\partial \Psi}{\partial r_2}\Big|_{t=t_c}, \quad \lambda_3(t_c) = 0, \quad H(t_c) = 0, \tag{2.23}$$

where $\nu \in \mathbb{R}$. Since the Hamiltonian has no explicit dependency on time, it follows that $H(t) = 0$ for all $t \in [0, t_c]$. Note that since the terminal condition is the same as in Equation 2.10, its derivatives are implicit from Equation 2.14. From Pontryagin's minimum principle, it follows that

$$-\lambda_1 v \sin\theta - \lambda_2 v \sin(\psi - \theta) - \lambda_3 \left[\frac{v}{r_2} \cos(\psi - \theta) + \frac{v}{r_1} \cos\theta \right] = 0. \tag{2.24}$$

Theorem 2. *For the PP time-optimal control problem, given the initial positions of P_1 and E, if P_2 initially lies in the non-degenerate region ($\mathcal{N}$), then the optimal control strategy of the evader involves simultaneous capture.*

Proof. Consider the case when $r_1(t_c) = \epsilon < r_2(t_c)$. This implies $\lambda_1(t_c) = \nu$, and $\lambda_2(t_c) = 0$. Note that the adjoint equations are linear in the co-states λ_2, λ_3, and since $\lambda_2(t_c) = 0$, the co-states are constant in time, i.e., $\lambda_1(t) = \nu$, $\lambda_2(t) = 0$, $\lambda_3(t) = 0$. From Equation 2.22 and Equation 2.24, it follows that $-1 + \nu(-u_1 + v\cos\theta) = 0$ and $\nu v \sin\theta = 0$. If $\nu = 0$, the two equations will lead to a contradiction. Therefore, $\sin\theta = 0$. Thus $\theta^*(t) = 0$, which means that the optimal strategy is pure evasion from P_1. However, this is true only when the problem is degenerate. In a non-degenerate case, this would lead to an early capture by P_2. In this case, when $r_2(t_c) = \epsilon < r_1(t_c)$, we have $\lambda_1(t_c) = 0$ and $\lambda_2(t_c) = \nu$. Furthermore, from Equation 2.22 and Equation 2.24, at $t = t_c$ it follows that $-1 + \nu(-u_1 - v\cos(\psi(t_c) - \theta(t_c))) = 0$, and $\nu v \sin(\psi(t_c) - \theta(t_c)) = 0$. Since $\nu \neq 0$, it follows that $\sin(\psi(t_c) - \theta(t_c)) = 0$. With this terminal condition, it can be seen that the co-states are constant and so is the optimal heading, which in this case is given by $\theta^*(t) = \pi + \psi$. This means that the optimal evading strategy is a pure evasion from P_2. However, this strategy is infeasible in the non-degenerate case. Hence, the optimal evading strategy in a

non-degenerate case would result in $r_1(t_c) = r_2(t_c) = \epsilon$, i.e., a simultaneous capture. □

Proposition 1. *Given the initial positions of P_1 and E, the optimal control strategy of the evader in the case of PP can be summarized as follows:*

$$\tan\theta^* = \begin{cases} 0, & (x_2(0), y_2(0)) \in \mathcal{D}_2, \\ \Theta(r_1, r_2, \psi, \lambda_1, \lambda_2, \lambda_3), & (x_2(0), y_2(0)) \in \mathcal{N}, \\ \pi + \psi, & (x_2(0), y_2(0)) \in \mathcal{D}_1, \end{cases} \tag{2.25}$$

where $(x_2(0), y_2(0))$ is the initial position of P_2 in inertial frame, and

$$\Theta(r_1, r_2, \psi, \lambda_1, \lambda_2, \lambda_3) = -\frac{\lambda_2 \sin\psi + (\lambda_3/r_1) + (\lambda_3 \cos\psi/r_2)}{\lambda_1 - \lambda_2 \cos\psi + (\lambda_3 \sin\psi/r_2)}. \tag{2.26}$$

Proof. The proof is a direct consequence of defining the degenerate and non-degenerate regions for PP (see subsection 2.3.2). □

With this analysis, the optimal control problem for the case PP can be solved numerically. An analytical solution to the equations of optimality subject to arbitrary initial conditions is at this point elusive. Numerical results for the cases CB and PP are shown in subsection 2.4.4. Next, we present a suboptimal strategy for PP that is easy to implement in practice. This suboptimal strategy of the evader is based on geometric arguments and is discussed in the next subsection.

2.4.3 A Suboptimal Strategy for Pure Pursuit

The optimal strategy for PP can be intuitively understood as one where the evader chooses its heading so that it does not favor any one of the two pursuers, finally resulting in simultaneous capture by both pursuers. With this motivation, a suboptimal strategy is constructed and its performance is compared with the optimal one in the case of identical pursuers. In this regard, the time-to-capture relation given in Equation 2.5 is exercised. For a non-

degenerate problem, the evader's heading for which both P_1 and P_2 take equal time to reach the evader can be found from their initial positions using the expression

$$\frac{r_1(0)(u_1 + v\cos\theta)}{u_1^2 - v^2} = \frac{r_2(0)(u_2 + v\cos\theta)}{u_2^2 - v^2}. \tag{2.27}$$

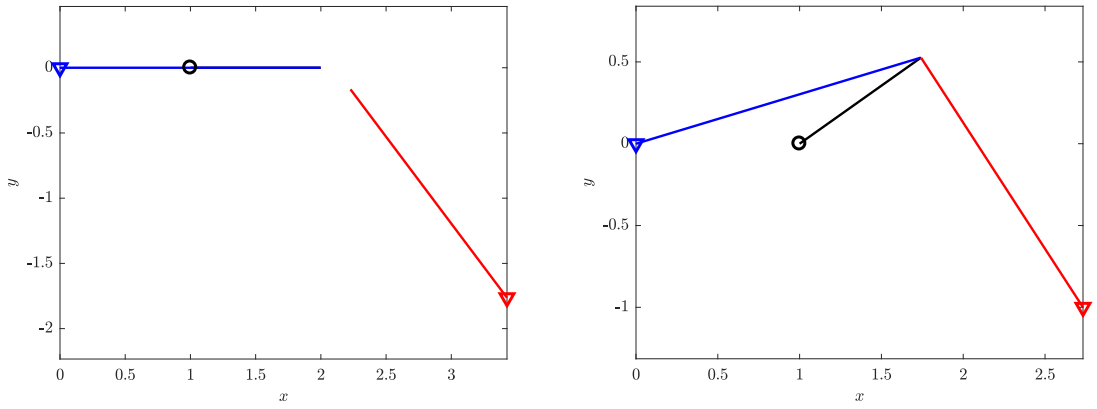
In general, Equation 2.27 has two solutions resulting in simultaneous capture, assuming the evader follows a constant heading. For the given initial conditions, the solution to Equation 2.27 that provides maximum capture time is chosen as the suboptimal strategy. If the problem is degenerate (with respect to P_1 or P_2), then Equation 2.27 has no solution.

2.4.4 Numerical Simulations

This subsection demonstrates the aforementioned optimal evading strategies using simulations performed for the cases CB and PP with different initial conditions. For simplicity, we assume that the speeds of the pursuers are the same and set to $u_1 = u_2 = 1$, whereas the speed of the evader is set to $v = 0.5$, unless specified otherwise. The radius of capture is chosen as $\epsilon = 0.001$.

The optimal strategy for CB is straightforward. The software package GPOPS-II [170] was used to simulate the test cases and validate the presented theory. Figure 2.9(a) presents the trajectories of the players for the initial conditions, $p_1 = (0, 0)$, $p_2 = (3.427, -1.763)$, $p_E = (1, 0)$, that makes the problem degenerate. Clearly, the optimal strategy is a pure evasion from P_1, and P_2 does not affect the evader's trajectory. An example for the non-degenerate case is presented in Figure 2.9(a) for the initial conditions, $p_1 = (0, 0)$, $p_2 = (2.732, -1)$, $p_E = (1, 0)$. It can be observed that the optimal evading strategy involves simultaneous capture with constant heading.

The simulation results for a non-degenerate case of PP, obtained using GPOPS-II, can be seen in Figure 2.10. Figure 2.10(a) presents the trajectories of the players for initial conditions $p_1 = (0, 0)$, $p_2 = (2.516, -0.875)$, $p_E = (1, 0)$. In the reduced state-space,



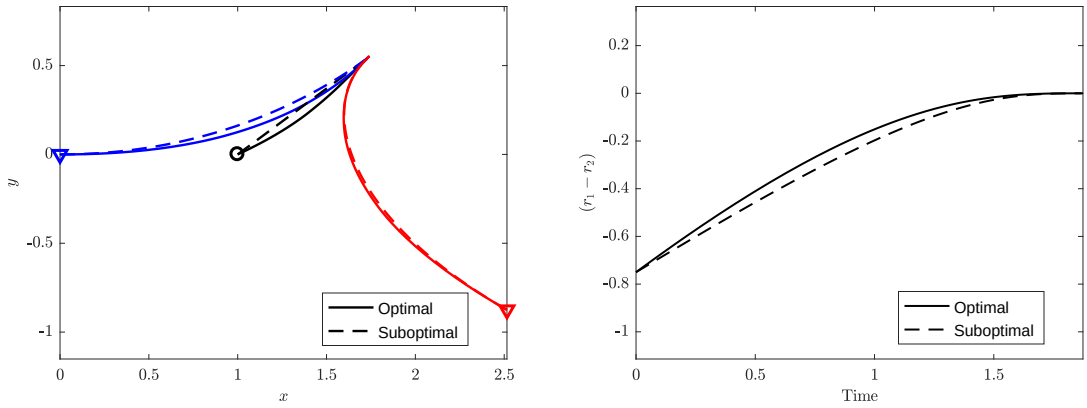
(a) A degenerate case

(b) A non-degenerate case

Figure 2.9: Trajectories of the players for optimal control inputs in the case of CB: black - evader, blue - P_1, red - P_2.

these positions correspond to $r_1(0) = 1$, $r_2(0) = 1.75$, and $\psi(0) = -\pi/6$. The optimal capture time is $t_c = 1.874$. The difference between the relative distances $(r_1 - r_2)$ is shown in Figure 2.10(b).

As expected, simultaneous capture is observed in these figures. Also, the difference in the relative distances, $(r_1 - r_2)$, becomes zero only at the final time. This suggests that the evader is equidistant from both the pursuers just before it gets captured. The same behavior has been observed in all the simulations that were carried out. The suboptimal strategy is also compared against the optimal strategy in Figure 2.10. The (constant) heading obtained from the suboptimal strategy is $\theta = 0.6378$ (36.54^o) with a capture time of $t_c = 1.868$. Note that the capture time and the variation in $(r_1 - r_2)$ are comparable to the corresponding results obtained using the optimal strategy, see Figure 2.10(b). Furthermore, a comparative study was carried out to gauge the performance of this suboptimal strategy. For this purpose, the following parameters were chosen: $r_1(0) = 1$, $u_1 = u_2 = u = 1$. The speed of the evader v was varied from 0.3 to 0.7. For each v, 140 different initial conditions $(r_2(0), \psi(0))$ were considered spanning the non-degenerate area for the chosen $r_1(0)$ and u. Table 2.1 presents the results of this comparative study. Though the average percentage variation of the time-to-capture increases with the evader's speed v, the variation is less



(a) Trajectories (b) Difference between the relative distances

Figure 2.10: Performance of the optimal and suboptimal strategies for a non-degenerate case of PP: black - evader, blue - P_1, red - P_2.

Table 2.1: Comparison table for optimal and suboptimal strategies for the case of PP

v	Average percentage variation in t_c	Maximum percentage variation in t_c
0.3	0.0337 %	0.4451 %
0.4	0.0727 %	0.7653 %
0.5	0.1277 %	1.2182 %
0.6	0.1704 %	1.6883 %
0.7	0.2343 %	2.2487 %

than 1% for all the evader speeds considered. The maximum percentage variation is only 2%. It can be observed that the suboptimal strategy is easily implementable and its performance is similar to the optimal one. Hence, the suboptimal strategy can be considered for all practical purposes.

2.5 Optimal Evading Strategies with a Stationary Pursuer

In this section, and without loss of generality, both pursuers are assumed to be identical, and one of the pursuers remains stationary during the game. This scenario may result, for instance, from the implementation of a relay-pursuit strategy, according to which only one pursuer is assigned to go after the evader at every instant of time [20, 40]. The pursuer whose Voronoi cell contains the evader is assigned to be the *active* pursuer to chase the evader. The other pursuer, designated as the *inactive* pursuer, stays at its original location

and plays the role of a guard. The active pursuer switches when the evader enters the interior of the Voronoi cell of another pursuer. Due to the symmetry of the problem, when the evader resides on the Voronoi boundary, we can assign any one of the two pursuers to be the active pursuer. Therefore, throughout the pursuit process, we can fix one of the pursuers to be the active pursuer, while the other pursuer remains stationary whose mere presence imposes a state restriction, namely that the evader does not enter the interior of its corresponding Voronoi cell.

Without loss of generality, we choose P_1 to be the active pursuer having velocity $u_1 = u$, and P_2 to be the inactive pursuer with velocity $u_2 = 0$, assumed to be located at the origin, i.e., $p_2(0) = (0, 0)$. The equations of motion can be obtained from Equation 2.1 with $u_1 = u$ and $u_2 = 0$. The game evolves in the four-dimensional state space, $[x_1, y_1, x_E, y_E]^\top \in \mathbb{R}^4$. We consider scenarios differing in the information structure and pursuit strategies used, subject to the state constraint

$$\|p_1 - p_E\| \leq \|p_2 - p_E\| = \|p_E\|. \tag{2.28}$$

This constraint restricts the evader from entering the Voronoi cell of the inactive pursuer. First, the region of non-degeneracy and the value of employing two pursuers in a relay pursuit mode is examined based on the set of initial conditions.

2.5.1 Region of Non-degeneracy

The problem is non-degenerate for a given set of initial conditions if the inactive pursuer affects the outcome of the game, i.e., if the optimal evading strategy is not pure evasion from the active pursuer. Therefore, for a given set of initial conditions, and with the evader following a pure evasion strategy from the active pursuer, if the evader enters the Voronoi section of the inactive pursuer before it gets captured, then the problem is non-degenerate and vice versa. Since the position of the inactive pursuer is fixed at the origin, given the

initial position of the active pursuer, the region of non-degeneracy is defined as the set of evader's initial positions for which the problem is non-degenerate. Note that the region of non-degeneracy is the same for both R-CB and R-PP, given the active pursuer's initial position, unlike for the case of two active pursuers analyzed in section 2.3. This is because of the fact that P_1 and P_2 are identical in terms of their speed capabilities and P_2 is stationary in a relay pursuit problem.

Proposition 2. *[20] In the two-pursuer one-evader problem with one inactive pursuer, the evader will be captured before entering the Voronoi cell of the inactive pursuer while moving along the LoS and away from the active pursuer, if and only if the quadratic equation*

$$at^2 + bt + c = 0, \tag{2.29}$$

where $a = u^2 - 2uv$, $b = 2[(ux_1(0) - vx_1(0) - ux_E(0))\cos\theta_E(0) + (uy_1(0) - vy_1(0) - uy_E(0))\sin\theta(0)]$ *and* $c = x_1(0)^2 + y_1(0)^2 - 2(x_E(0)x_1(0) + y_E(0)y_1(0))$*, does not have a solution inside the interval* $[0, t_f]$*, where* t_f *is obtained from Equation 2.2, and* $\theta_E(0)$ *is determined by the equations:*

$$\cos\theta_E(0) = \frac{x_E(0) - x_1(0)}{\|p_E(0) - p_1(0)\|}, \qquad \sin\theta_E(0) = \frac{y_E(0) - y_1(0)}{\|p_E(0) - p_1(0)\|}. \tag{2.30}$$

Notice that when $v < u/2$, there exists no initial position for the evader such that the condition in Proposition 2 is satisfied. Therefore, the optimal control for the evader is always to move along the LoS of P_1 when $v < u/2$. Henceforth, we assume that $v > u/2$.

In order to find the explicit expression for the region satisfying the condition in Proposition 2, without loss of generality, let the initial position of the active pursuer be $p_1 = (x_1(0), 0)$. A schematic for the non-degenerate region with $p_1 = (-2, 0)$, $v = 0.8$, $u = 1$ is shown in Figure 2.11. The green region depicts the evader's initial positions for which the condition in Proposition 2 is not satisfied. That is, if the evader starts from a position inside the shaded region, it will not be able to move along the LoS of P_1 throughout the

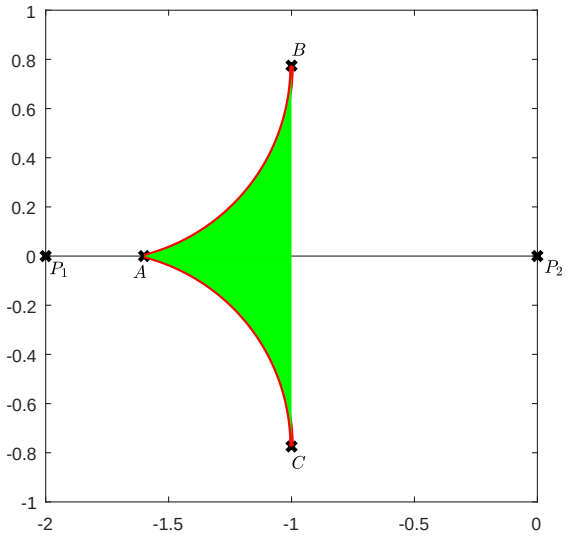


Figure 2.11: Region of non-degeneracy

pursuit without violating the state constraint. We denote the three vertices of the shaded region by A, B and C, where A resides on the line segment between the active pursuer and the inactive pursuer, and B, C are on the Voronoi boundary.

If the evader starts on the line segment P_1P_2 at an initial position $(x_E(0), 0)$ that does not violate the state constraint, while moving along the LoS of P_1, it must be captured by the P_1 before it reaches the boundary of the Voronoi cell. Thus, $(|x_1(0)| - |x_E(0)|)/(u - v) \leq (|x_E(0)| - |x_1(0)|/2)/(v - u/2)$. After simplification, we have $|x_E(0)| \geq |x_1(0)|v/u$. Since A is at the boundary of the region and the previous inequality is linear with respect to $x_E(0)$, we obtain $A = (x_1(0)v/u, 0)$.

Since B is the uppermost point of the shaded region that is also on the Voronoi boundary, when the evader starts at B, its velocity for staying on the Voronoi boundary should coincide with its velocity for moving along the LoS. On the boundary, we have $x_1(x_1 - 2x_E) + y_1(y_1 - 2y_E) = 0$. Hence, by taking a time derivative, we obtain $-2(u\|p_E - p_1\| - vx_1 \cos\theta_E - vy_1 \sin\theta_E) = 0$. By plugging in $x_E(0) = x_1(0)/2, y_E(0) = \beta, y_1(0) = 0$ in

the previous equations, one obtains

$$\cos\theta_{_E}(0) = -\frac{u(x_1(0)^2/4+\beta^2)^{1/2}}{vx_1(0)}. \tag{2.31}$$

On the other hand, for the evader to move along the LoS of P_1, $\theta_{_E}(0)$ satisfies

$$\cos\theta_{_E}(0) = \frac{-x_1(0)/2}{(x_1(0)^2/4+\beta^2)^{1/2}}. \tag{2.32}$$

Solving for β using the expressions in Equation 2.31 and Equation 2.32, we obtain $B = \left(x_1(0)/2, \sqrt{v/(2u)-1/4}\,|x_1(0)|\right)$. And, $C = \left(x_1(0)/2, -\sqrt{v/(2u)-1/4}\,|x_1(0)|\right)$, which by the nature of symmetry is the reflection of B about the x-axis.

The curves AB and AC are arcs of circles and satisfy the equation

$$0 = (((v-u)x_1(0)+ux)(x-x_1(0))+uy^2)^2-u(2v-u)x_1(0)(2x-x_1(0))((x-x_1(0))^2+y^2), \tag{2.33}$$

which is derived from $0 = b^2 - 4ac$, where a, b and c are defined in Proposition 2. The optimal evading strategies can now be analyzed in the non-degenerate regions for R-CB and R-PP.

2.5.2 Constant Bearing Strategy

In the case of R-CB, it is assumed that the active pursuer (P_1) follows a constant bearing strategy. The reduced state space dynamics are given by

$$\dot{r}_1 = v\cos(\theta_{_E}-\varphi_{10}) - u\cos(\theta_1-\varphi_{10}), \tag{2.34a}$$

$$\dot{r}_2 = v\cos(\theta_{_E}-\varphi_2), \tag{2.34b}$$

$$\dot{\varphi}_2 = \frac{v}{r_2}\sin(\theta_{_E}-\varphi_2), \tag{2.34c}$$

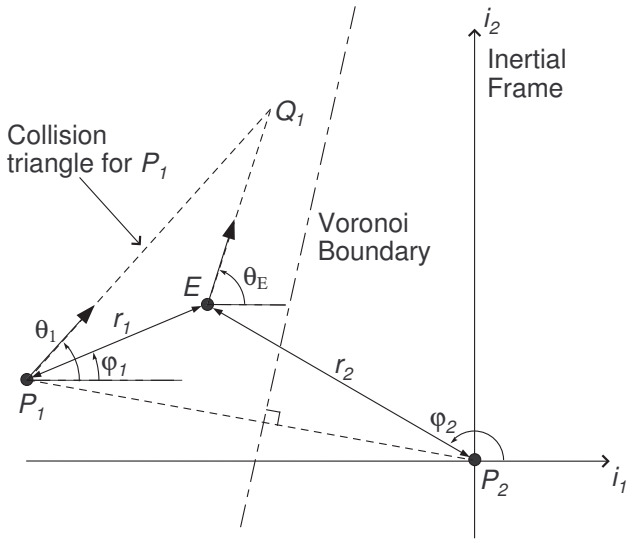


Figure 2.12: Schematic of the proposed relay pursuit problem with the active pursuer following a constant bearing strategy (R-CB)

where φ_{10} is the initial LoS angle of P_1. Since P_1 follows a constant bearing strategy, θ_1 is determined from the equation

$$\dot{\varphi}_1 = \frac{1}{r_1}\left[v\sin(\theta_{_E} - \varphi_{10}) - u\sin(\theta_1 - \varphi_{10})\right] = 0. \tag{2.35}$$

These equations are similar to the ones presented in subsection 2.4.1, but they differ in the sense that the second pursuer is stationary.

A schematic of the proposed pursuit-evasion problem can be seen in Figure 2.12. The boundary conditions are given by $r_1(0) = \|p_{_E}(0) - p_1(0)\|, r_2(0) = \|p_{_E}(0) - p_2(0)\|, \varphi_2(0) = \varphi_{20}, r_1(t_c) = \epsilon$, $r_2(t_c)$ and $\varphi_2(t_c)$ are free, where φ_{20} is the initial LoS angle of the inactive pursuer (P_2). The state constraint in Equation 2.28, imposed on the relay pursuit problem, can be expressed as

$$\mathcal{S} = \frac{1}{2}(r_1^2 - r_2^2) \leq 0. \tag{2.36}$$

Note that the inequality constraint in Equation 2.36 involves only the state variables. Therefore, we have to take the time derivative for the constraint $\mathcal{S}$ and substitute the equa-

tions of motion until an explicit dependence on the control variable occurs. The qth-order time derivative in which this first happens plays a role in the Hamiltonian [165]. To this end, we first have

$$\dot{\mathcal{S}} = r_1\dot{r}_1 - r_2\dot{r}_2 = r_1(v\cos(\theta_{_E} - \varphi_{10}) - u\cos(\theta_1 - \varphi_{10})) \; - vr_2\cos(\theta_{_E} - \varphi_2). \quad (2.37)$$

It can be seen that the control variable θ appears explicitly in the first time derivative. The Hamiltonian can then be expressed as

$$\begin{aligned} H(r_1, r_2, \psi_2, \lambda_1, \lambda_2, \lambda_3, \mu) &= -1 + \lambda_1\dot{r}_1 + \lambda_2\dot{r}_2 + \lambda_3\dot{\varphi}_2 + \mu\dot{\mathcal{S}} \\ &= -1 + (\lambda_1 + \mu r_1)(v\cos(\theta_{_E} - \varphi_{10}) - u\cos(\theta_1 - \varphi_{10})) \\ &\quad + (\lambda_2 - \mu r_2)v\cos(\theta_{_E} - \varphi_2) + \frac{\lambda_3 v}{r_2}\sin(\theta_{_E} - \varphi_2), \quad (2.38) \end{aligned}$$

where λ_1, λ_2, λ_3, and μ are the co-states, μ satisfies the Kuhn-Tucker and complementary slackness conditions i.e., for $\mathcal{S} \neq 0$, $\mu = 0$, and $\mathcal{S} = 0$, $\mu \geq 0$. The transversality conditions are given by $\lambda_2(t_c) = 0, \lambda_3(t_c) = 0, H(t_c) = 0$. Furthermore, since the Hamiltonian has no explicit dependency on time, $H(t) = 0$ for all $t \in [0, t_c]$. The optimal control can be obtained from Pontryagin's minimum principle, using the expression

$$(\lambda_1+\mu r_1)\sin(\theta_1-\theta_{_E})-(\lambda_2-\mu r_2)\sin(\theta_{_E}-\varphi_2)\cos(\theta_1-\varphi_{10})+\frac{\lambda_3}{r_2}\cos(\theta_{_E}-\varphi_2) = 0. \quad (2.39)$$

Proposition 3. *In the case of R-CB, if the initial conditions are such that the problem is*

non-degenerate with $u > v > u/2$, then the optimal control of the evader $\theta_E^*(t)$ satisfies

$$\begin{cases} \lambda_1 \sin(\theta_1(t) - \theta_E^*(t)) - \lambda_2 \sin(\theta_E^*(t) - \varphi_2) \cos(\theta_1(t) - \varphi_{10}) & \\ +\lambda_3 \cos(\theta_E^*(t) - \varphi_2)/r_2 = 0, & t \in [0, \tau_1], \\ r_1(v \cos(\theta_E^*(t) - \varphi_{10}) - u \cos(\theta_1(t) - \varphi_{10})) & \\ -v r_2 \cos(\theta_E^*(t) - \varphi_2) = 0, & t \in [\tau_1, \tau_2], \\ \theta_E^*(t) = \varphi_{10}, & t \in [\tau_2, t_c], \end{cases}$$

where $\theta_1(t)$ is obtained using the expression

$$v \sin(\theta_E^*(t) - \varphi_{10}) = u \sin(\theta_1(t) - \varphi_{10}), \quad t \in [0, \tau_2] \tag{2.40}$$

Furthermore, τ_2 satisfies the switching condition:

$$(v - u) r_1(\tau_2) - v r_2(\tau_2) \cos(\varphi_{10} - \varphi_2(\tau_2)) = 0. \tag{2.41}$$

2.5.3 Pure Pursuit Strategy

In this case, P_1 follows a pure pursuit strategy. A schematic of the proposed relay pursuit problem is shown in Figure 2.13. This problem is similar to the one presented in subsection 2.4.2, and can be analyzed in the three-dimensional reduced state space, but it differs from it by the fact that P_2 is now stationary. It can be solved with the use of states and the control input θ in the reduced state space presented in subsection 2.4.2.

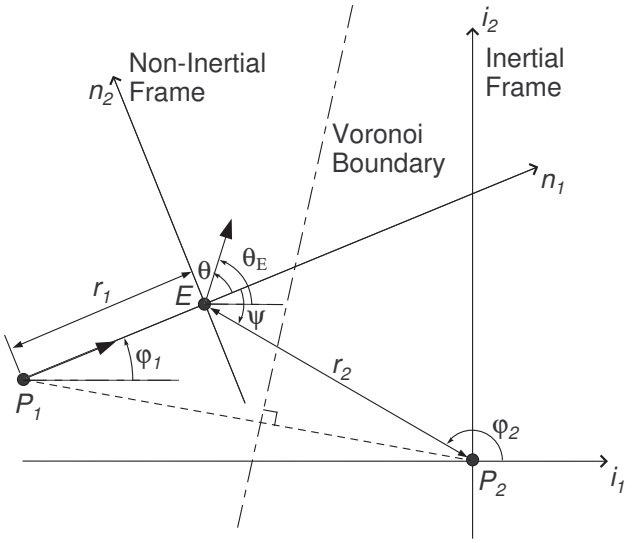


Figure 2.13: Schematic of the proposed relay pursuit problem with the active pursuer following a pure pursuit strategy (R-PP)

The equations of motion are now given by

$$\dot{r}_1 = -u + v\cos\theta, \tag{2.42a}$$

$$\dot{r}_2 = -v\cos(\psi - \theta), \tag{2.42b}$$

$$\dot{\psi} = \frac{v}{r_2}\sin(\psi - \theta) - \frac{v}{r_1}\sin\theta, \tag{2.42c}$$

while the constraint given in Equation 2.36, and the boundary conditions for r_1 and r_2 are the same as in the R-CB case. Note that $\psi(0) = \psi_0$ and $\psi(t_c)$ is free. Also, a time derivative of the constraint $\mathcal{S}$ has to be taken in order to write the Hamiltonian for this optimal control problem and

$$\dot{\mathcal{S}} = r_1\dot{r}_1 - r_2\dot{r}_2 = r_1(-u + v\cos\theta) + vr_2\cos(\psi - \theta). \tag{2.43}$$

Therefore, the Hamiltonian is

$$H(r_1, r_2, \psi, \lambda_1, \lambda_2, \lambda_3, \mu) = -1 + (\lambda_1 + \mu r_1)(-u + v\cos\theta) + (-\lambda_2 + \mu r_2)v\cos(\psi - \theta) + \lambda_3(\frac{v}{r_2}\sin(\psi - \theta) - \frac{v}{r_1}\sin\theta), \tag{2.44}$$

where λ_1, λ_2, λ_3, and μ are the co-states, μ satisfies the Kuhn-Tucker and complementary slackness conditions i.e., for $\mathcal{S} \neq 0$, $\mu = 0$, and $\mathcal{S} = 0$, $\mu \geq 0$. The transversality conditions are given by $\lambda_2(t_c) = 0, \lambda_3(t_c) = 0$ and $H(t_c) = 0$. Again, the Hamiltonian has no explicit dependency on time and therefore, the optimal Hamiltonian is zero for all time. The optimal control can be obtained from the Pontryagin's minimum principle, using the expression

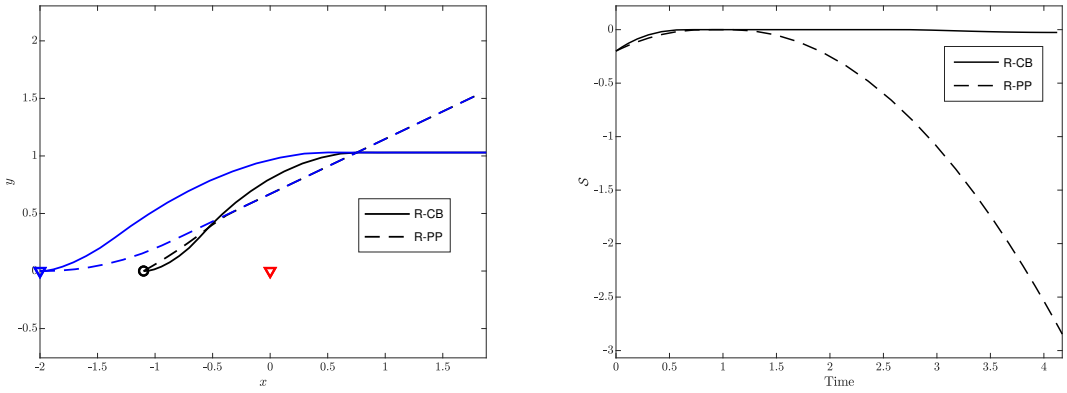
$$\tan\theta^* = -\frac{(\lambda_2 - \mu r_2)\sin\psi + (\lambda_3/r_1) + (\lambda_3\cos\psi/r_2)}{\lambda_1 - (\lambda_2 - \mu r_2)\cos\psi + (\lambda_3\sin\psi/r_2)}. \tag{2.45}$$

Proposition 4. *[20] In the case of R-PP, if the initial conditions are such that the problem is non-degenerate with $u > v > u/2$, then the optimal control of the evader $\theta^*(t)$ satisfies*

$$\tan\theta^*(t) = \begin{cases} -\dfrac{\lambda_2\sin\psi + (\lambda_3/r_1) + (\lambda_3\cos\psi/r_2)}{\lambda_1 - \lambda_2\cos\psi + (\lambda_3\sin\psi/r_2)}, & t \in [0, \tau_1], \\ \dfrac{q - \sigma p\sqrt{p^2 + q^2 - 1}}{p + \sigma q\sqrt{p^2 + q^2 - 1}}, & t \in [\tau_1, \tau_2], \\ 0, & t \in [\tau_2, t_c], \end{cases}$$

where $p = (vr_1 + vr_2\cos\psi)/(ur_1)$, $q = (vr_2\sin\psi)/(ur_1)$, $\sigma = \mathrm{sgn}(q)$. Furthermore, τ_2 satisfies the switching condition:

$$(v - u)r_1(\tau_2) + vr_2(\tau_2)\cos\psi(\tau_2) = 0. \tag{2.46}$$



(a) Trajectories (b) State constraint ($\mathcal{S}$)

Figure 2.14: Performance of the optimal strategies for a non-degenerate cases in R-CB and R-PP: black - evader, blue - P_1, red - P_2

Summarizing the previous analyses, we conclude that in both cases R-CB and R-PP, the optimal trajectory of the evader involves three periods. First the evader moves inside the Voronoi cell of the active pursuer in a way such that the optimal conditions (corresponding transversality conditions, Erdmann corner conditions, etc.) are satisfied before it hits the Voronoi boundary. The evader then moves along the boundary until the switching condition: Equation 2.41 for R-CB; Equation 2.46 for R-PP; is satisfied. Finally, the evader moves along the LoS of P_1 till capture occurs.

2.5.4 Numerical Simulations

In this subsection the optimal strategies for R-CB and R-PP are demonstrated using an example with non-degenerate initial conditions. The speed of active pursuer (P_1) is $u = 1$ and that of evader is $v = 0.8$. The initial position of P_1 is $(-2, 0)$ and evader's initial position is at $(-1.1, 0)$ which is in the non-degenerate region, see Figure 2.11. Since the non-degenerate region is the same in both cases, their corresponding optimal strategies can be compared in this example. The problem is solved using GPOPS-II.

The simulated results are presented in Figure 2.14. The time-to-capture in the case of R-PP is $t_c = 4.172$, while in R-CB, $t_c = 4.125$. The R-CB case has a lower capture

time when compared to R-PP because in the latter, the pursuers have an advantage as they have information of both the evader's position and velocity. The trajectories of the players can be seen in Figure 2.14(a). Since P_1 follows a constant bearing strategy, its LoS angle is constant throughout the time ($\varphi_{10} = 0$). The state constraint can be analyzed using Figure 2.14(b). The three periods in both optimal strategies can be observed in that plot. The interesting observation here is that in R-CB the evader hits the Voronoi boundary earlier compared to R-PP, but the evader stays on the boundary for longer time and eventually follows pure evasion from P_1. This is because of the difference in P_1's strategy, which also affects the dynamics of the Voronoi boundary, and the switching condition.

CHAPTER 3

TASK ALLOCATION IN MULTI-PLAYER PURSUIT-EVASION PROBLEMS

This chapter addresses multi-pursuer multi-evader (MPME) problems by developing scalable solutions based on a decomposition of the original MPME problem to a sequence of simpler multi-pursuer single-evader (MPSE) problems. A major enabler for this decomposition is a new result that allows us to characterize each pursuer as relevant or redundant for each evader. Only the relevant pursuers participate in the MPSE pursuit of each evader. The identification and classification of each pursuer as relevant or redundant makes use of the classical tool of the Apollonius circle (in the case when the pursuers follows a constant bearing strategy) or its extension, termed herein as Apollonius curve, (in the case when the pursuers follows a pure pursuit strategy). The efficacy of the approach is demonstrated using an illustrative example involving 10 pursuers and 5 evaders.

3.1 Problem Formulation

In this chapter, we consider a pursuit-evasion problem in the Euclidean plane that involves n identical pursuers and m identical evaders. The pursuers' objective is to capture all the evaders. Capture occurs when one or more pursuers enter the capture zone of an evader (assumed here to be a disk of radius $\epsilon > 0$ centered at the instantaneous position of the evader). At the same time, each evader aims at avoiding capture indefinitely. Let P_i denote the i^{th} pursuer and let E_j denote the j^{th} evader. Let also $\mathcal{P} = \{P_1, P_2, \ldots, P_n\}$ denote the set of pursuers and, similarly, let $\mathcal{E} = \{E_1, E_2, \ldots, E_m\}$ denote the set of evaders. With a slight abuse of notation, in the sequel we will also use the subscript indices to denote the corresponding evader or pursuer.

The equations of motion of all the agents are given below

$$\dot{x}_i = u\cos\theta_i, \quad \dot{y}_i = u\sin\theta_i, \quad i \in \mathcal{P}, \tag{3.1a}$$

$$\dot{x}_j = v\cos\theta_j, \quad \dot{y}_j = v\sin\theta_j, \quad j \in \mathcal{E}, \tag{3.1b}$$

where $p_i = (x_i, y_i) \in \mathbb{R}^2$, and $e_j = (x_j, y_j) \in \mathbb{R}^2$ denote the positions of pursuer P_i, and evader E_j, respectively, and θ_i and θ_j denote the heading angles (control inputs) for the pursuers and the evaders, respectively. In Equation 3.1, u and v are the speeds of the pursuers and the evaders, which are assumed constant with $u > v$. The number of states is $2(n + m)$, and it increases linearly with the number of players. It is assumed that the pursuers follow a given, known pursuit strategy. As discussed in chapter 2, in this chapter too, two distinct pursuit strategies are investigated: CB - the pursuers follow a constant bearing strategy; and PP - the pursuers follow a pure pursuit strategy.

The three problems to be addressed in this chapter are listed below.

Problem 1: For both cases CB and PP with $m = 1$ (MPSE problem), and assuming that the pursuers are unaware of the evader's strategy, which pursuers should go after the evader to minimize capture time?

Problem 2: For both cases CB and PP with $m = 1$ (MPSE problem), and assuming that the evader has complete information of the pursuers' whereabouts and their strategy, what is the time-optimal evading strategy?

Problem 3: For both cases CB and PP with $m \geq 1$, and assuming that the pursuers are unaware of the evaders' strategy, which pursuer(s) should go after which evader(s)?

3.2 Optimal Evading Strategies in Multi-Pursuer Single-Evader Problems

3.2.1 Constant Bearing Strategy

A schematic of the problem geometry with one evader and n pursuers (henceforth referred to as the MPSE problem) following a constant bearing strategy is shown in Figure 3.1.

Since the pursuers are assumed to be following a constant bearing (CB) strategy, the problem can be analyzed by tracking the relative distances between the pursuers and the evader, effectively reducing the number of states from $2(n+1)$ to just n. In this regard, the dynamics can be written in the form

$$\dot{r}_i = v\cos(\theta_{_E} - \varphi_i) - u\cos(\theta_i - \varphi_i), \qquad i \in \mathcal{P}, \tag{3.2}$$

where r_i is the relative distance between pursuer P_i and the evader, and $\varphi_i = \text{atan2}(y_{_E} - y_i, x_{_E} - x_i)$ is the corresponding line of sight (LoS) angle. From now on we will drop the subscripts for the evader and will use E instead of j to denote the single evader in MPSE settings (i.e., in section 3.2 and section 3.3). Furthermore, we indicate the pursuers using the subscripts directly and the set $\mathcal{P} = \{1, 2, \ldots, n\}$. Note that in the case of a constant bearing strategy, the bearing angle between a pursuer and the evader remains constant until the time of capture. Using this fact, the instantaneous heading of pursuer P_i (θ_i, $i \in \mathcal{P}$) can be obtained from the relation,

$$u\sin(\theta_i - \varphi_i) = v\sin(\theta_{_E} - \varphi_i), \tag{3.3}$$

which is a function of the instantaneous heading of the evader $\theta_{_E}$. The above relation has two possible solutions for each θ_i, given $\theta_{_E}$, and the solution for which $\dot{r}_i < 0$ is chosen.

The initial conditions of the problem are

$$r_i(0) = \|e(0) - p_i(0)\|, \qquad i \in \mathcal{P}, \tag{3.4}$$

and the terminal condition is

$$\Psi(r_1(t_c), r_2(t_c), \ldots, r_n(t_c)) = \min_{i\in\mathcal{P}} r_i(t_c) - \epsilon = 0, \tag{3.5}$$

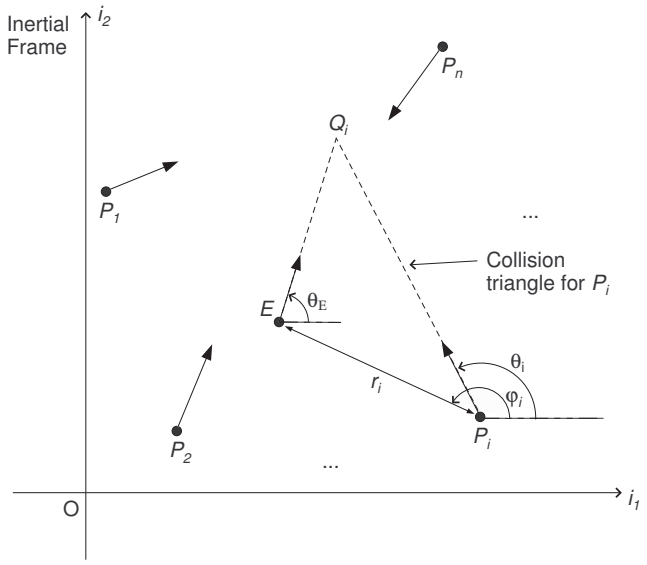


Figure 3.1: Schematic of the proposed MPSE problem with pursuers following a constant bearing strategy

where t_c is the time-of-capture. A formal definition for the time-of-capture is given below.

Definition 1. *Given the initial positions of the players (at $t = 0$) in an n-pursuer single-evader problem, and assuming that the pursuers follow either a constant bearing or a pure pursuit strategy, and for any evading strategy,* the time-to-capture t_c (≥ 0) *is the minimum time so that there is at least one pursuer in the capture zone of the evader.*

Since θ_i is a function of θ_E, and by the nature of Assumption A1 ($v < u$), r_i decreases monotonically for all time, and lies in the set $[\epsilon, r_i(0)]$, $i \in \mathcal{P}$. Also, the time of capture is finite for all evader strategies, and is bounded by $t_c^{\max}$, which is the capture time for the farthest pursuer (at the initial time among the n pursuers), assuming the evader follows a pure evasion strategy and none of the pursuers move. Note that the dynamics of the evader is a function of just the control $\theta_E \in [0, 2\pi]$, and therefore $\dot{r}_i \in [-(v+u), v-u]$, for all $i \in \mathcal{P}$, at any time and state, which is a convex set. Therefore, from Filippov's theorem [171], there exists a time-optimal evading strategy. This fact and the necessary conditions given below together ensure that the proposed evading strategy is indeed optimal.

To continue with our analysis, note that since the evader strives to maximize t_c using its

control input θ_E, the Hamiltonian of the underlying optimal control problem can be written as

$$H = -1 + \sum_{i \in \mathcal{P}} \lambda_i \left[v \cos(\theta_E - \varphi_i) - u \cos(\theta_i - \varphi_i) \right], \tag{3.6}$$

where $\lambda_i \, (i \in \mathcal{P})$ are the costates. The corresponding adjoint equations are obtained as

$$\dot{\lambda}_i = -\frac{\partial H}{\partial r_i} = 0, \qquad i \in \mathcal{P}, \tag{3.7}$$

and hence the costates are constants, $\lambda_i(t) = c_i$, $t \in [0, t_c]$, for all $i \in \mathcal{P}$. The transversality conditions are

$$\lambda_i(t_c) = \nu \frac{\partial \Psi}{\partial r_i}\bigg|_{t=t_c}, \quad i \in \mathcal{P}, \quad \text{and} \quad H(t_c) = 0. \tag{3.8}$$

Since the Hamiltonian has no explicit dependency on time, it follows that

$$H(t) = 0, \quad t \in [0, t_c]. \tag{3.9}$$

Note that the terminal condition is not differentiable and its partial with respect to each component $r_i(t_c)$ can be expressed as

$$\begin{aligned}
&\frac{\partial \Psi}{\partial r_i}\bigg|_{t=t_c} = 0, && \text{if } r_i(t_c) > \epsilon, \\
&\frac{\partial \Psi}{\partial r_i}\bigg|_{t=t_c} = 1, && \text{if } r_i(t_c) = \epsilon \text{ and } r_j(t_c) \neq \epsilon, \quad j \neq i, \\
&\frac{\partial \Psi}{\partial r_i}\bigg|_{t=t_c} \text{ is undefined}, && \text{if } r_i(t_c) = \epsilon \text{ and } r_j(t_c) = \epsilon, \text{ for some } j \neq i.
\end{aligned} \tag{3.10}$$

Applying the Potryagin's Minimum Principle, with $\dfrac{\partial H}{\partial \theta_{_E}} = 0$, yields

$$\sum_{i \in \mathcal{P}} \lambda_i \left[-v \sin(\theta_{_E} - \varphi_i) + u \sin(\theta_i - \varphi_i) \frac{\partial \theta_i}{\partial \theta_{_E}} \right] = 0, \tag{3.11}$$

where from Equation 3.3,

$$\frac{\partial \theta_i}{\partial \theta_{_E}} = \frac{v \cos(\theta_{_E} - \varphi_i)}{u \cos(\theta_i - \varphi_i)}, \qquad \cos(\theta_i - \varphi_i) \neq 0. \tag{3.12}$$

Now, the following definition is used to establish the characteristics of the optimal evading strategy.

Definition 2. *Consider an MPSE problem and assume that the pursuers follow either a constant bearing or a pure pursuit strategy. For a given strategy of the evader,* the capturing pursuer set $\mathcal{P}_c \subset \mathcal{P}$ *is the set of pursuers that are in the capture zone of the evader at* t_c.

Refer to Definition 1 for t_c (time-to-capture), which is always finite since the pursuers follow either a constant bearing or a pure pursuit strategy. Note that at the time of capture, one or more pursuers can be in the capture zone of the evader. Therefore, $1 \leq \text{card}[\mathcal{P}] \leq n$, where card$[\cdot]$ represents the cardinality of the set. The capturing pursuer set for the optimal evading strategy, given the pursuers follow either a constant bearing or a pure pursuit strategy, is denoted by $\mathcal{P}^*$

Proposition 5. *In the case of an MPSE problem with all the pursuers following a constant bearing strategy, the time-optimal evading strategy is dependent only on the initial positions of those pursuers that are in the corresponding capturing pursuer set* $\mathcal{P}^*$.

Proof. Let card$[\mathcal{P}^*] = k$, $1 \leq k \leq n$ and, without loss of generality, assume that pursuers $P_1, P_2, \ldots, P_k$, capture the evader simultaneously at time t_c. Therefore, $r_1(t_c) = \cdots = r_k(t_c) = \epsilon$. Then, from Equation 3.10, $\lambda_i(t) = c_i = 0$, for $i \in \mathcal{P} \backslash \mathcal{P}^*$, and Equation 3.11

can be written as

$$\sum_{i\in\mathcal{P}^*} \lambda_i \left[-v\sin(\theta_{_E} - \varphi_i) + u\sin(\theta_i - \varphi_i)\frac{\partial\theta_i}{\partial\theta_{_E}} \right] = 0. \tag{3.13}$$

Since $\theta_i\,(i \in \mathcal{P}^*)$ is a function of $\theta_{_E}$, and the LoS angles φ_i are dependent only on the initial positions of the players, from Equation 3.13, it is evident that the optimal control input of the evader is only dependent on the initial positions of those pursuers that capture it at the final time t_c. □

Substituting Equation 3.12 in Equation 3.13, the later can be simplified to

$$\sum_{i\in\mathcal{P}^*} \lambda_i \left[\frac{\sin(\theta_i - \theta_{_E})}{\cos(\theta_i - \varphi_i)} \right] = 0. \tag{3.14}$$

The above equation cannot be simplified any further to obtain a closed form optimal strategy for the evader and hence, the information about the set of pursuers that capture the evader at the time-of-capture i.e., the set $\mathcal{P}^*$, cannot be obtained in an analytic fashion. Numerical examples indicate that the problem may contain multiple local minima. To tackle this problem and to address the issue of pursuer allocation, the idea of active/redundant pursuers is introduced in the next section. Following the discussion on optimal strategies, its characteristics for the case of pure pursuit is discussed in the following subsection.

3.2.2 Pure Pursuit Strategy

In the case of a pure pursuit strategy, the velocity vector of the pursuer is aligned along the LoS, as shown in Figure 3.2 i.e., the LoS angles (φ_i) do not remain constant anymore. Therefore, the dynamics has to include the evolution of both relative distances and the

corresponding LoS angles, which can be written as

$$\dot{r}_i = -u + v\cos(\theta_E - \varphi_i), \tag{3.15}$$

$$\dot{\varphi}_i = \frac{v}{r_i}\sin(\theta_E - \varphi_i), \qquad i \in \mathcal{P}. \tag{3.16}$$

The initial conditions include Equation 3.4 and

$$\varphi_i(0) = \text{atan2}(y_E(0) - y_i(0), x_E(0) - x_i(0)), \qquad i \in \mathcal{P}, \tag{3.17}$$

with the terminal condition being the same as in Equation 3.5.

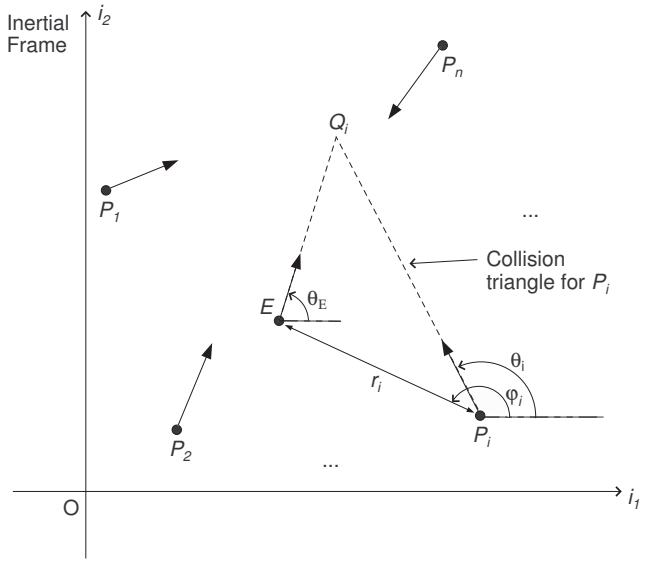


Figure 3.2: Schematic of the proposed MPSE problem with pursuers following a pure pursuit strategy

For the case of an MPSE problem where all pursuers follow a pure pursuit strategy, contrary to the constant bearing case of subsection 3.2.1, it is not easy to show existence of an optimal evading strategy using Filippov's theorem (although a feasible evading strategy always exists trivially). Hence, in the following discussion we make the implicit assumption that a time-optimal evading strategy exists, and we proceed to characterize this strategy using the necessary conditions for optimality.

The Hamiltonian of the time-optimal control problem in the case of PP can be written as

$$H = -1 + \sum_{i \in \mathcal{P}} \left[\lambda_i \left(-u + v \cos(\theta_E - \varphi_i) \right) + \mu_i \frac{v}{r_i} \sin(\theta_E - \varphi_i) \right], \tag{3.18}$$

and the adjoint equations are

$$\dot{\lambda}_i = -\frac{\partial H}{\partial r_i} = -\mu_i \frac{v}{r_i^2} \sin(\theta_E - \theta_i), \tag{3.19}$$

$$\dot{\mu}_i = -\frac{\partial H}{\partial \varphi_i} = \lambda_i v \sin(\theta_E - \theta_i) - \mu_i \frac{v}{r_i^2} \cos(\theta_E - \theta_i), \quad i \in \mathcal{P}. \tag{3.20}$$

The transversality conditions include Equation 3.8 and

$$\mu_i(t_c) = \nu \frac{\partial \Psi}{\partial \varphi_i}\bigg|_{t=t_c} = 0, \qquad i \in \mathcal{P}. \tag{3.21}$$

Furthermore, since the Hamiltonian has no explicit dependence on time, it is zero for the entire time interval and Equation 3.9 holds for this case as well. The partials with respect to r_i are given in Equation 3.10. Finally, we have

$$\sum_{i \in \mathcal{P}} \left[-\lambda_i \sin(\theta_E - \varphi_i) + \frac{\mu_i}{r_i} \cos(\theta_E - \varphi_i) \right] = 0. \tag{3.22}$$

Proposition 6. *In the case of an MPSE problem with all the pursuers following a pure pursuit strategy, the time-optimal evading strategy is dependent only on the initial positions of those pursuers that are in the corresponding capturing pursuer set $\mathcal{P}^*$.*

Proof. Let card$[\mathcal{P}^*] = k$, $1 \leq k \leq n$, without loss of generality, assume that pursuers $P_1, P_2, \ldots, P_k$ capture the evader simultaneously at time t_c. Therefore, $r_1(t_c) = \cdots = r_k(t_c) = \epsilon$. Then, from Equation 3.10, $\lambda_i(t) = c_i = 0$, for $i \in \mathcal{P} \backslash \mathcal{P}^*$. Furthermore, from Equation 3.10, $\lambda_i(t_c) = 0$, for $i \in \mathcal{P} \backslash \mathcal{P}^*$, and from Equation 3.21, $\mu_i(t_c) = 0$, for $i \in \mathcal{P} \backslash \mathcal{P}^*$. Note that the adjoint equations, Equation 3.19 and Equation 3.20, for

all $i \in \mathcal{P}$ are affine in their respective costates. Therefore, for $i \in \mathcal{P}\backslash\mathcal{P}^*$, $\lambda_i(t) = 0$, $\mu_i(t) = 0,\ t \in [0, t_c]$. Hence, Equation 3.22 can be rewritten as

$$\sum_{i\in\mathcal{P}^*}\left[-\lambda_i \sin(\theta_{_E} - \varphi_i) + \frac{\mu_i}{r_i}\cos(\theta_{_E} - \varphi_i)\right] = 0. \tag{3.23}$$

In Equation 3.23, φ_i is dependent only on its initial conditions and the strategy of the evader. Clearly, the optimal control input of the evader is only dependent on the initial positions of those pursuers that capture it at the final time t_c. □

The pure pursuit case does not allow for a closed-form solution of the optimal evading strategy, and as a result, it is difficult to the obtain the set $\mathcal{P}^*$ analytically. In this regard, the following section presents suboptimal solutions for allocating the pursuers for the MPSE problem that can be employed under any evading strategy, while guaranteeing capture.

3.3 Active/Redundant Pursuers

This section presents strategies for the task allocation problem using the tool of Apollonius circles [10]. The Apollonius circle for a pursuer-evader pair is the locus of points where capture occurs, for all possible initial headings of a non-maneuvering evader, given the initial positions of the pursuer-evader pair and assuming that the pursuer follows a constant bearing strategy, see Figure 3.3. For the MPSE problem, the Apollonius circle of the pair P_i-E is denoted as $\mathcal{A}_i$. It has its center at $O_i\left(\frac{x_{_E} - \rho x_i}{1-\rho^2}, \frac{y_{_E} - \rho y_i}{1-\rho^2}\right)$ and radius $d_i = \frac{\rho}{1-\rho^2}\|p_i - e\|$, where $\rho = v/u$ (speed ratio) [32]. The Apollonius circles evolve in time as the players move, but the time dependencies will be dropped for the sake of brevity. Let T_i be the closest point to the evader on the Apollonius circle where collision occurs when the evader goes head-on with the pursuer, as shown in Figure 3.3. Therefore the distance of T_i from the evader is $v\|p_i - e\|/(u + v)$.

When the pursuer employs a pure pursuit strategy, the locus of all points where capture occurs is a closed curve, represented using $\mathcal{C}_i$ (see Figure 3.3), and designated as an

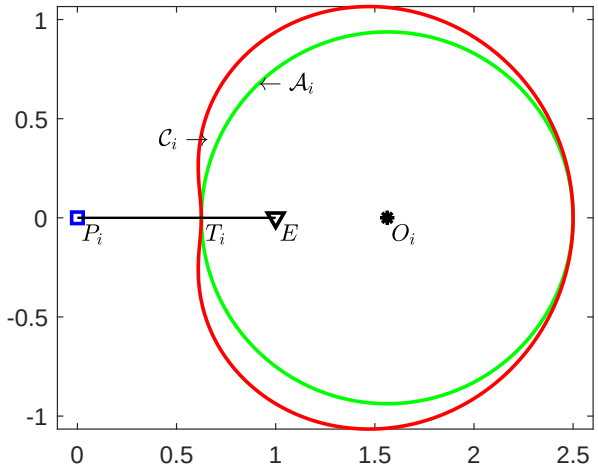


Figure 3.3: The locus of capture points for a non-maneuvering evader in the cases CB and PP (simulation parameters: $u = 1$, $v = 0.6$, $p_i(0) = (0, 0)$, $p_E = (1, 0)$)

Apollonius curve. This Apollonius curve can be obtained from the time taken to capture a non-maneuvering evader

$$t_f = \frac{r_o(u + v\cos\phi_i)}{u^2 - v^2}, \qquad v \neq u, \tag{3.24}$$

where r_o is the initial distance between the pursuer-evader pair, and ϕ is the evader's heading measured with respect to the line-of-sight from the pursuer to the evader at the initial time (see Figure 3.2) [167]. Furthermore, given the heading of a non-maneuvering evader, the t_f-isochrone in the case of a pure pursuit strategy always contains the t_f-isochrone of a constant bearing strategy [167, 172]. Therefore, the time-to-capture in the later case is either higher than or equal to the former case. This gives rise to the following lemma.

Lemma 1. *Given the positions of the pursuer P_i and the evader E, the corresponding Apollonius circle $\mathcal{A}_i$ is always contained in the area enclosed by the Apollonius curve $\mathcal{C}_i$.*

The Apollonius curve is the locus of capture points for a non-maneuvering evader when the pursuer uses a pure pursuit strategy. The Apollonius curve thus generalizes the notion of the Apollonius circle for the case of pure pursuit. Note that the area enclosed by the curve

$\mathcal{C}_i$ forms a non-convex set and hence, $\mathcal{C}_i$ is a non-convex curve [173]. Next, the Apollonius circle $\mathcal{A}_i$ and the Apollonius curve $\mathcal{C}_i$ will be used to identify the *active* and *redundant* pursuers. The following definitions establish the notions of active and redundant pursuers. Please refer to Definition 2 for $\mathcal{P}_c$ (capturing pursuer set).

Definition 3. *Consider an MPSE problem and assume that all pursuers follow either a constant bearing or a pure pursuit strategy. If $P_i \in \mathcal{P}_c$ for some evading strategy, then P_i is an* active pursuer. *Otherwise, P_i is a* redundant pursuer.

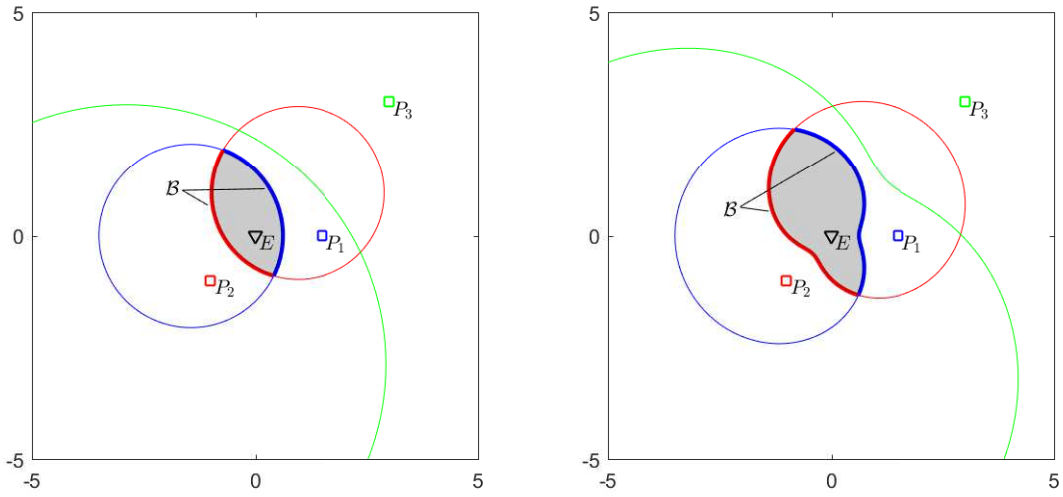
Given the instantaneous positions of the pursuers and the evader, it is of interest to find a condition to verify whether a pursuer is active or redundant. In this regard, we first define the instantaneous *Apollonius boundary*.

Definition 4. *Given the positions of the players in an n-pursuer single-evader problem at time $0 \leq t < t_c$, and assuming that the pursuers follow a constant bearing strategy, the* Apollonius boundary *at time t is the set of points $\mathcal{B}_t = \{X \in \bigcup_{i=1}^{n} \mathcal{A}_i \mid \mathcal{M}(e, X) \cap \left(\bigcup_{i=1}^{n} \mathcal{A}_i\right) = \{X\}\}$, where $\mathcal{M}(e, X)$ denotes the set of points on the line segment with endpoints e (position of the evader) and X at time t.*

In other words, the Apollonius boundary is the set of points that belong to the union of all the instantaneous Apollonius circles, and in addition, each such point is the closest to the evader along its respective line-of-sight originating from the evader. The following lemma establishes an important property of the Apollonius boundary, which will be used in section 3.4.

Lemma 2. *Given the positions of the players in an n-pursuer single-evader problem at time $0 \leq t < t_c$, and assuming that the pursuers follow a constant bearing strategy, the Apollonius circle of the closest pursuer is always a part of the Apollonius boundary $\mathcal{B}_t$.*

Proof. Without loss of generality, assume P_1 is the closest pursuer. It follows that $\underset{i \in \mathcal{P}}{\operatorname{argmin}} \|p_i - e\| = 1$, and point T_1 (the point closest to the evader on the Apollonius circle $\mathcal{A}_1$, see



(a) CB (b) PP

Figure 3.4: Apollonius circles, curves and boundaries for CB and PP cases (Simulation parameters: $u = 1$, $v = 0.6$)

Figure 3.3) is the closest point to the evader along the corresponding line-of-sight originating from the evader. Therefore, the point T_1 satisfies the condition $\mathcal{M}(e, T_1) \cap \left(\bigcup_{i=1}^{n} \mathcal{A}_i \right) = \{T_1\}$. Hence, the Apollonius circle of the closest pursuer is always a part of the Apollonius boundary. □

Similarly, the Apollonius boundary in the case of pursuers following a pure pursuit strategy can be defined with $\mathcal{C}_i$ replacing $\mathcal{A}_i$ in Definition 4. The Apollonius boundaries in both cases can be visualized in Figure 3.4. It can be observed that the region enclosed by the Apollonius boundary in the case of CB is always convex but it may not be the case for PP. Note that the Apollonius boundary evolves with time as the Apollonius circles or curves evolve with time as well.

3.3.1 Identifying Active/Redundant Pursuers

An algorithm to identify active/redundant pursuers in the case of CB is discussed first, which is based on the following conjecture.

Conjecture 1. *Given the positions of all the players in an MPSE problem at time $0 \leq t < t_c$, and assuming that the pursuers follow a constant bearing strategy, pursuer P_i is active at time t if $\mathcal{B}_t \cap \mathcal{A}_i \neq \varnothing$, and is redundant otherwise.*

The conjecture implies that in the case of CB, a pursuer is active at time $0 \leq t < t_c$, if and only if its corresponding Apollonius circle is part of the Apollonius boundary at that instant. The conjecture is inspired from the fact that the region in which the capture point lies in is bounded by the instantaneous Apollonius circle for any strategy of the evader. Note that if a pursuer is active at time t', it need not remain active for all $t > t'$. But if a pursuer is redundant at time t', it will remain redundant for all $t > t'$. The following lemmas based on Conjecture 1 provide simple checks to determine whether a pursuer is active or redundant.

Lemma 3. *Given the positions of the players in an n-pursuer single-evader problem at time $0 \leq t < t_c$, assume that the pursuers follow a constant bearing strategy. Then pursuer P_i is the only active pursuer if and only if the conditions*

$$\mathcal{A}_i \bigcap \left(\bigcup_{j=1, j\neq i}^{n} \mathcal{A}_j \right) = \varnothing, \tag{3.25}$$

$$\mathcal{M}(e, T_i) \bigcap \left(\bigcup_{j=1, j\neq i}^{n} \mathcal{A}_j \right) = \varnothing, \tag{3.26}$$

are satisfied, where T_i is the closest point to the evader on the Apollonius circle $\mathcal{A}_i$. Furthermore, if conditions in Equation 3.25 and Equation 3.26 are not satisfied, then P_i is a redundant pursuer.

Proof. The necessity of the conditions in Equation 3.25 and Equation 3.26 for P_i to be the only active pursuer is established first. From Equation 3.25, it can be seen that $\mathcal{A}_i$ does not intersect any other Apollonius circle. Therefore, it is either the case that $\mathcal{A}_i$ contains all other Apollonius circles, or $\mathcal{A}_i$ is contained in every other Apollonius circle. Note that both cases are mutually exclusive. In the later case, the Apollonius boundary is $\mathcal{A}_i$ itself,

and P_i is the only active pursuer. Now, if

$$\mathcal{M}(e, X) \bigcap \left(\bigcup_{j=1,\, j\neq i}^{n} \mathcal{A}_j \right) = \varnothing, \tag{3.27}$$

where X is any point on $\mathcal{A}_i$, then $\mathcal{A}_i$ is contained in every other Apollonius circle, that is, P_i is the only active pursuer. Since X can be any point on the Apollonius circle $\mathcal{A}_i$, a convenient way to check the condition in Equation 3.27 is to choose the closest point to the evader on the Apollonius circle $\mathcal{A}_i$. This point has the closed form expression $T_i\left(\frac{x_E + \rho x_i}{1+\rho}, \frac{y_E + \rho y_i}{1+\rho}\right)$, see Figure 3.3. Conversely, if P_i is the only active pursuer, then from Conjecture 1, the Apollonius boundary is $\mathcal{A}_i$ itself. In such a case, $\mathcal{A}_i$ does not intersect any other Apollonius circle and it is contained in every other Apollonius circle, which implies Equation 3.25 and Equation 3.26 hold. Thus, Equation 3.25 and Equation 3.26 become necessary and sufficient conditions for P_i to be the only active pursuer. □

Lemma 4. *Given the positions of the players in an MPSE problem at time $0 \leq t < t_c$, assume that the pursuers follow a constant bearing strategy, and that the Apollonius circle $\mathcal{A}_i$ intersects at least one of the other Apollonius circles. Then, pursuer P_i is an active pursuer if and only if there exists $X \in \mathcal{I}_i$ such that*

$$\mathcal{M}(e, X) \bigcap \left(\bigcup_{j=1}^{n} \mathcal{A}_j \right) = \{X\}, \tag{3.28}$$

where $\mathcal{I}_i$ is the set of intersection points between $\mathcal{A}_i$ and the rest of the Apollonius circles.

Proof. The necessity of condition Equation 3.28 for P_i to be an active pursuer is proven first. Note that $X \in \mathcal{I}_i \subset \mathcal{A}_i$. Since $X \in \mathcal{B}_t$, and from Conjecture 1, it follows that P_i is an active pursuer at time t. Conversely, if P_i is an active pursuer at time t, from Lemma 4, and since $\mathcal{A}_i$ intersects one or more Apollonius circles, $\mathcal{A}_i$ alone cannot form the Apollonius boundary (see Lemma 3). Therefore, only portion(s) of $\mathcal{A}_i$ (i.e., arc(s) of the circle) can be a part of the Apollonius boundary. The arc(s) which could possibly be a part of $\mathcal{B}_t$

will have one or more of the intersection points as its endpoints. Hence, if P_i is an active pursuer, then there is at least one intersection point $X \in \mathcal{I}_i$ that satisfies the condition in Equation 3.28. □

The set of intersection points $\mathcal{I}_i$ can be obtained analytically given the instantaneous positions of all the players [174]. The above two lemmas can be used to verify if a pursuer is active or redundant. In this regard, Algorithm 1 below, named Apollonius circle based Active Pursuer Check (AAPC), can be employed to check the status of each pursuer. The time complexity of the algorithm is of order $O(n^2)$, since the maximum number of intersections between any two circles is two. Note that by dynamically allocating the task of capturing the evader using AAPC (where at every instant the active pursuers keep pursuing the evader while the redundant pursuers do not react), the pursuers as a group will be able to capture the evader in minimum time. Furthermore, if a pursuer becomes redundant at any point of time $0 \leq t < t_c$, it remains redundant after that (i.e., till capture occurs).

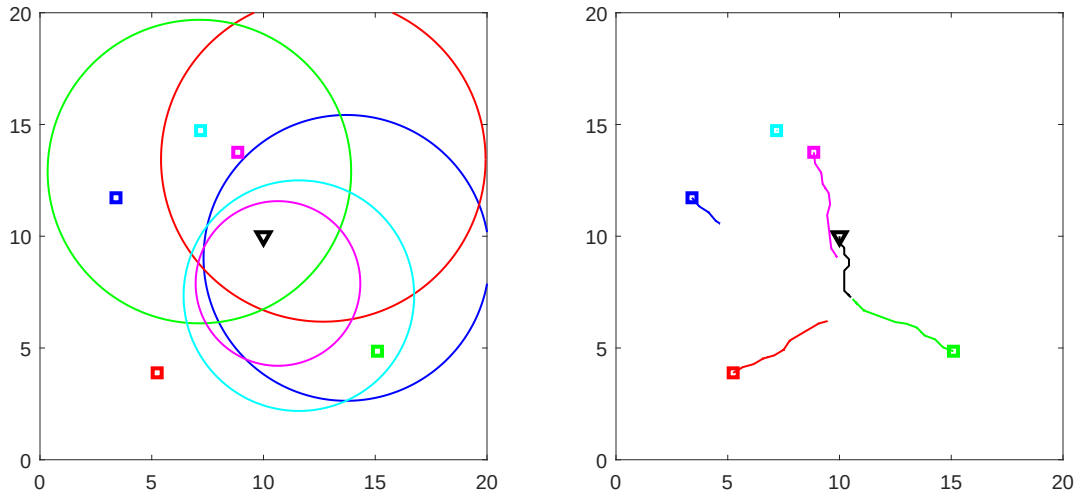
The case of PP is more involved because the corresponding Apollonius curve is non-convex. A claim similar to the one given in Conjecture 1 cannot be made and hence, it is difficult to determine the status of a pursuer in this case. In this regard, the convex hull of the area surrounded by the Apollonius curve can be considered. The boundary of this convex hull is used to obtain a refined Apollonius curve, and the active/redundant pursuers can be identified by having checks similar to the ones given in the case of CB. In this case, the active pursuers are simply the ones that keep pursuing the evader. The redundant ones are the ones that remain at rest. At the same time, and since the refined Apollonius curve in the case of PP does not have any closed form expression, obtaining the intersection points between the refined curves or between the refined curves and a given line segment is computationally more involved. Using an algorithm analogous to AAPC with refined Apollonius curves, numerical simulations obtained in the case of PP are presented in the following subsection. Note that the refined Apollonius curve is a convex curve [173].

Remark 1. *The notion of regions of non-degeneracy (RND), discussed in chapter 2, is taken*

Algorithm 1 Apollonius circle based Active Pursuer Check (AAPC)

Require: Positions of all the players (p_1,...,p_n,e,i)
Ensure: Status of pursuer P_i

procedure OBTAIN_STATUS(p_1,...,p_n,e,i)
    flag1 = 0 *(To check if $\mathcal{A}_i$ intersects any other Apollonius circle)*
    status = redundant
    for $j = 1$ to n and $j \neq i$ **do**
        Obtain $\mathcal{I}_{ij}$ *(set of intersection points (X_ℓ) for $\mathcal{A}_i$ and $\mathcal{A}_j$)*
        if $\mathcal{I}_{ij} \neq \varnothing$ **then**
            flag1 = 1
            for $\ell = 1$ to card$[\mathcal{I}_{ij}]$ **do**
                flag2 = 0. *(To check if $\mathcal{M}(e, X_\ell)$ intersects any other Apollonius circle)*
                for $k = 1$ to n and $k \neq i,\ j$ **do**
                    if $\mathcal{M}(e, X_\ell)$ intersects $\mathcal{A}_k$ **then**
                        flag2 = 1
                    end if
                end for
                if flag2 = 0 **then**
                    status = active
                    break from outermost loop.
                end if
            end for
        end if
    end for
    if flag1 = 0 **then**
        status = active
        for $j = 1$ to n and $j \neq i$ **do**
            if $\mathcal{M}(e, T_i)$ intersects $\mathcal{A}_j$ **then**
                status = redundant
                break
            end if
        end for
    end if
    return status
end procedure



(a) Initial Apollonius circles

(b) Trajectories

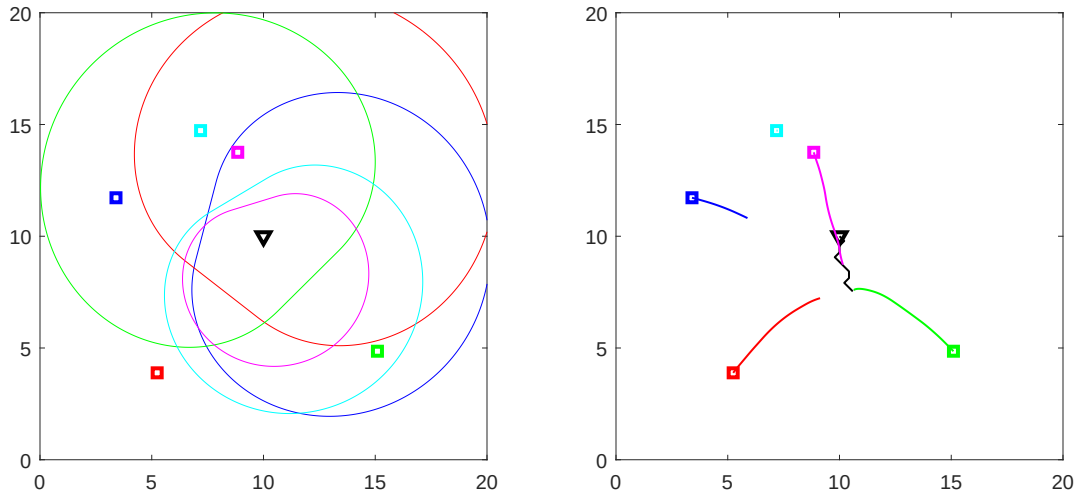
Figure 3.5: Results obtained using AAPC for task allocation in the case of CB

care of implicitly in the proposed approach using the notion of active/redundant pursuers. In chapter 2, RND were used to check whether a pursuer is active/redundant in two-pursuer one-evader problems. These were based on the location of the farther pursuer *compared to the* closer pursuer. *However, when there are more than two pursuers RND as an approach no longer works.*

3.3.2 Numerical Simulations

In this section, simulations of pursuer allocation using AAPC for both CB and PP cases, involving five pursuers and one evader, are presented. The speeds of the pursuers are set to $u = 1$, whereas the speed of the evader is set to $v = 0.6$. The radius of capture is chosen as $\epsilon = 0.1$. The evader follows a form of blind evasion strategy with switching times that are predefined [175]. At each switching time, the evader randomly chooses a heading from a set of allowable headings. The allowable headings set that is specific to the example showcased in this paper is $\{-\pi/4,\ \pi/2,\ 3\pi/4\}$.

Figure 3.5 presents the results obtained for the case of CB. Figure 3.5(a) shows the



(a) Initial Apollonius circles (refined) (b) Trajectories

Figure 3.6: Results obtained using AAPC with refined Apollonius curves for task allocation in the case of PP

initial positions of all the players along with the corresponding Apollonius circles. The triangle denotes the initial position of the evader and the square markers denote the initial positions of the pursuers. It can be observed that at the initial time, the pursuers identified with the colors red, magenta, green, and blue are the active pursuers, as their corresponding Apollonius circles are part of the Apollonius boundary. Figure 3.5(b) shows the trajectories of all the players. It can be seen that the green pursuer finally captures the evader, and the rest of the three pursuers, which are initially active, become redundant as time progresses. The cyan pursuer, which is redundant at the initial time, does not move at all. Figure 3.6 presents the results obtained for the PP case using the refined Apollonius curves. An analysis similar to its CB counterpart can be made by observing Figure 3.6(a) and Figure 3.6(b), where the cyan pursuer is again redundant at the initial time. The rest of the pursuers, though initially active, eventually become redundant except for the green pursuer, which finally captures the evader, see Figure 3.6(b).

3.4 Extension to Multi-Pursuer Multi-Evader Problems

In this section, the AAPC is extended to solve MPME problems. Given the positions of all the players at some instant of time, the set of evaders for which a pursuer is active can be obtained using AAPC. Note that at a given time instant, a pursuer can be classified as active by more than one evader or no evader whatsoever. In the case where a pursuer is classified as active for more than one evader, one can break the tie by assigning the pursuer to the nearest evader among the ones for which this pursuer is active. Using this idea, the following algorithm can be used for pursuer allocation in MPME problems. Note that the pursuers are assumed to be following a constant bearing strategy.

Apollonius-Voronoi Allocation Algorithm (AVAA): At a given time instant $0 \leq t \leq t_c$, let $\mathcal{E}_f$ be the set of evaders that are yet to be captured, and let $\mathcal{E}_c$ be the set of evaders that have already been captured. Note that $\mathcal{E} = \mathcal{E}_f \cup \mathcal{E}_c$. Given the current positions of all the players, let $\mathscr{I} : \mathcal{E}_f \to 2^{\mathcal{P}}$ be the initial allocation function that maps each evader E_j (in $\mathcal{E}_f$) to its set of active pursuers obtained by considering the positions of all the pursuers. That is, for a given $j \in \mathcal{E}_f$, $\mathscr{I}(j)$ is a subset of $\mathcal{P}$. Furthermore, $\mathcal{P}_a = \bigcup\limits_{j \in \mathcal{E}_f} \mathscr{I}(j)$ denotes the set of all the active (or assigned) pursuers according to the initial allocation function $\mathscr{I}$. Given the initial allocation function $\mathscr{I}$, let now $\mathscr{J} : \mathcal{P} \to 2^{\mathcal{E}_f}$ be the dual function defined by $\mathscr{J}(i) = \{j \in \mathcal{E}_f : \mathscr{I}(j) = i\}$. In other words, $\mathscr{J}$ maps each pursuer to the set of the evaders to which it is allotted as per $\mathscr{I}$. Next, we define the *final allocation function* $\mathscr{F}$ and the *intermediate allocation function* $\mathscr{G}$ as follows.

(a) If card$[\mathscr{J}(i)] \leq 1$, for all $i \in \mathcal{P}$, then let $\mathscr{F} = \mathscr{I}$. Otherwise, let $\mathscr{G} : \mathcal{E} \to 2^{\mathcal{P}}$ be defined as $\mathscr{G}(j) = \left\{ i \in \mathscr{I}(j) : j = \underset{k \in \mathscr{J}(i)}{\operatorname{argmin}} \|p_i - p_k\| \right\}$. The function $\mathscr{G}$ maps each evader to a set of pursuers in accordance to the mapping $\mathscr{I}$, such that each active pursuer is assigned to the nearest evader among its assigned ones. Note that $\mathscr{G}(j)$ can be an empty set for some j, i.e., an evader can end up be unassigned as per $\mathscr{G}$.

(b) Let $\mathcal{P}_u = \mathcal{P} \backslash \mathcal{P}_a$ be the set of unassigned pursuers. Now for each evader E_j, find

the active pursuers considering the positions of the pursuers that are only in the set $\mathscr{G}(j) \cup \mathcal{P}_u$, and obtain an *updated allocation function* $\mathscr{I}'$ and its corresponding dual $\mathscr{J}'$.

(c) Repeat steps (a) and (b), by replacing $\mathscr{I}$ and $\mathscr{J}$ with $\mathscr{I}'$ and $\mathscr{J}'$, respectively, until $\mathscr{F}$ is obtained.

Note that in step (a) of the algorithm, ties with multiple assignments of the same pursuer are broken using distance as the metric. Furthermore, if each pursuer is assigned to only one evader or if it remains unassigned, then the initial allocation function $\mathscr{I}$ is also the final one. In any other case, once the intermediate allocation function $\mathscr{G}$ is obtained in step (b) of the algorithm, the set of unassigned pursuers according to $\mathscr{I}$ is obtained. In step (b), an updated allocation function $\mathscr{I}'$ is obtained by checking for active pursuers among the set of unassigned pursuers coupled with the pursuers assigned as per $\mathscr{G}$, for each evader. Because one of the unassigned pursuers (in the set $\mathcal{P}_u$) can become active to the evaders that have lost one or more pursuers during the tie break in step (a). With $\mathscr{I}'$ and its corresponding dual $\mathscr{J}'$, steps (a) and (b) are repeated until each pursuer has only one (or none) assignment. Once an evader is captured, it is removed from the set $\mathcal{E}_f$ and added to the set $\mathcal{E}_c$.

The above algorithm is run at every time instant to obtain the allocation function $\mathscr{F}$, given the players' current positions, until all the evaders are captured, i.e., until $\mathcal{E}_f$ is empty. The algorithm provides a potentially sub-optimal solution, but it is scalable for any number of pursuers and evaders. The algorithm guarantees capture of all m evaders as is shown in Theorem 3 below. In order to prove this theorem, several preparatory results are needed.

Definition 5. *Given the positions of the players in an MPME problem at time $t \geq 0$, the* current shortest reach *(CSR) is defined by* $\min\limits_{(i,j)\in\mathcal{P}\times\mathcal{E}_f} \|p_i - e_j\|$.

Lemma 5. *At a given time instant $t \geq 0$, $i^* \subseteq \mathscr{F}(j^*)$, where $(i^*, j^*) = \operatorname{argmin}\limits_{(i,j)\in\mathcal{P}\times\mathcal{E}_f} \|p_i - e_j\|$, and $\mathscr{F}$ is the final allocation function of AVAA.*

Proof. From Lemma 2, since i^* is the closest pursuer to j^*, the Apollonius boundary for j^* contains i^*, while considering the positions of all the pursuers (in $\mathcal{P}$). Therefore, i^* will be assigned to j^*, as per the initial allocation function $\mathscr{I}$. When i^* has multiple assignments as per $\mathscr{I}$, the intermediate allocation function $\mathscr{G}$ still assigns i^* to j^*, as the pursuer is assigned to the closest evader in the case of multiple assignments (as per $\mathscr{G}$). Furthermore, i^* is assigned to j^*, as per all the subsequent updated allocation functions $\mathscr{I}'$, owing to the fact i^* is the closest pursuer to j^*, and hence as per the final allocation function $\mathscr{F}$. □

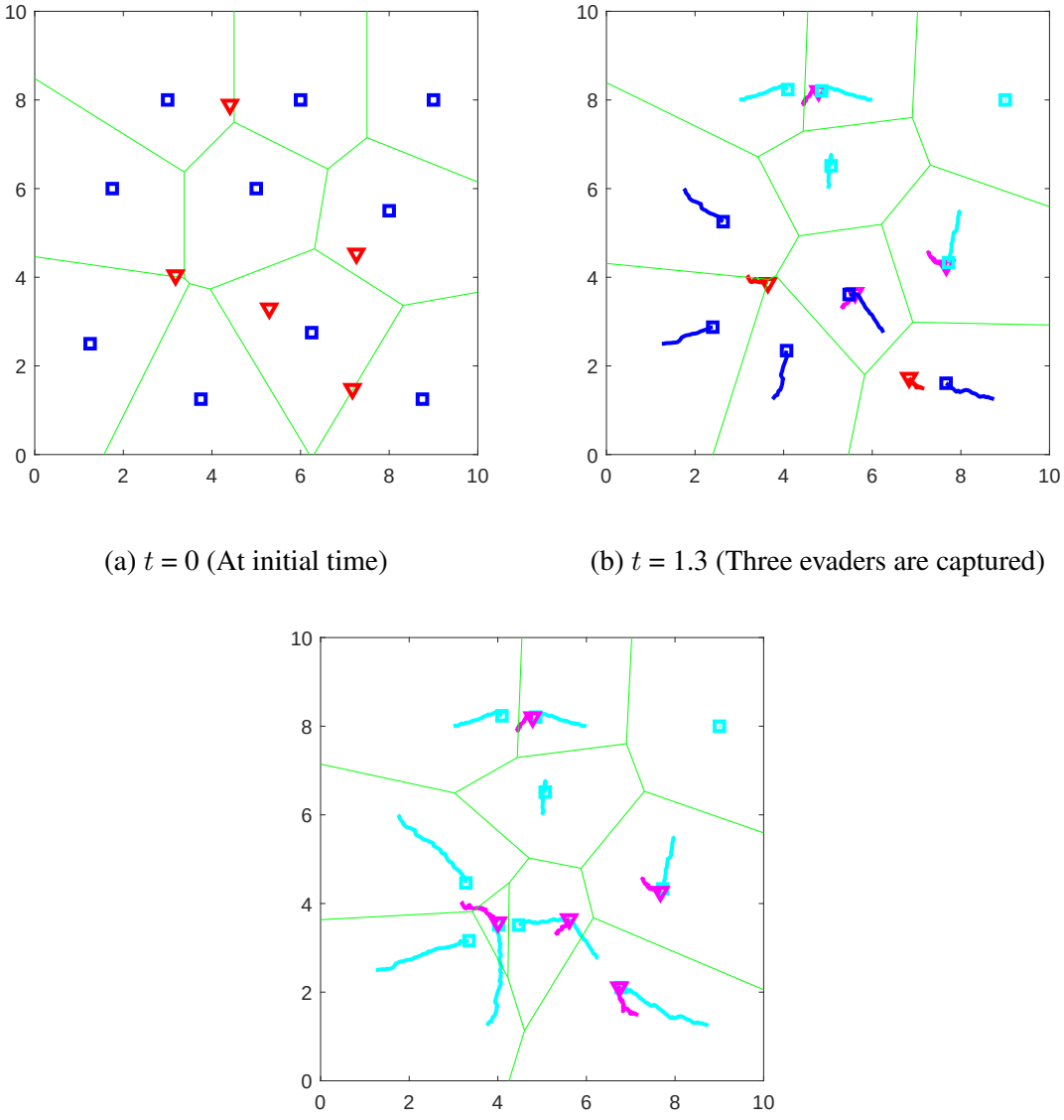
Lemma 6. *Assuming the pursuers are assigned to the evaders using AVAA, at any given time $t \geq 0$, CSR will converge to zero in finite time, and hence at least one evader will be captured in finite time.*

Proof. From Lemma 5, pursuer i^* (corresponding to the CSR) is always assigned to evader j^*. Since all the pursuers are faster compared to the evaders, and since they follow a constant bearing strategy, $\mathrm{d}(\|p_{i^*} - e_{j^*}\|)/\mathrm{dt} < 0$, for all $t \geq 0$ [167]. Furthermore, as the initial CSR is finite, the CSR converges to zero in finite time. Hence, capture of one evader is guaranteed in finite time. □

Theorem 3. *The AVAA algorithm guarantees capture of all the evaders in finite time.*

Proof. The result immediately follows from Lemma 5 and Lemma 6. Note that the CSR is updated (from zero) every time a capture occurs, and the captured evader is removed from the list of participating players. Also, the number of evaders are finite. □

Figure 3.7 demonstrates the performance of AVAA for 10 pursuers and 5 evaders. The simulation parameters remain the same as in subsection 3.3.2. In Figure 3.7, the red triangles indicate the current positions of the evaders that are not captured and the magenta ones are the evaders that are captured. The blue squares indicate the current positions of the active pursuers and the cyan ones indicate the redundant pursuers. In all three plots, the



(a) $t = 0$ (At initial time)

(b) $t = 1.3$ (Three evaders are captured)

(c) $t = 2.5$ (All evaders are captured)

Figure 3.7: Plots showing the positions and trajectories of the players in a multi-pursuer (squares) multi-evader (triangles) problem at different time instants

Voronoi partition of the domain with the pursuers as generators is also included for reference. The animation corresponding to the simulation shown in Figure 3.7, can be found on the web[1].

3.5 Pursuit-Evasion with Heterogeneous Teams

In the previous section, the proposed MPME formulation was analyzed with the pursuit and the evading teams being homogeneous i.e., all the pursuers are identical with equal speed capabilities, and so is the case with the evaders. In this section, assuming the pursuers follow a constant bearing strategy, we extend the results in section 3.4 to include heterogeneous teams of agents with guarantees on finite-time capture. To this end, u_i and v_j are the speeds of P_i, $i \in \mathcal{P}$, and E_j, $j \in \mathcal{E}$, which are assumed to be constant with

$$\min_{i \in \mathcal{P}} u_i > \max_{j \in \mathcal{E}} v_j. \tag{3.29}$$

The definition of the Apollonius circle The definitions for active and redundant pursuers in the case of homogeneous agent groups (Definition 3) can be directly adopted in the case of heterogeneous agent groups.

In the case where a pursuer is classified as active for more than one evader, one can break the tie by assigning the pursuer to the evader that can be captured in minimum possible time i.e., $\mathrm{argmin}_j \|p_i - e_j\|/(u_i + v_j)$. The aforementioned criterion for breaking a tie is equivalent to choosing the nearest evader when the teams are assumed to be homogeneous. The resulting Apollonius allocation (A2) algorithm, obtained by updating the tie-breaking criterion in the AVAA, can be used for pursuer allocation in MPME problems involving heterogeneous teams. The proof for finite-time capture of all evaders is presented below.

Definition 6. *Given the instantaneous positions of the players in the MPME problem at time $t \geq 0$, the* current shortest time *(CST) is defined by* $t_s = \min_{(i,j) \in \mathcal{P} \times \mathcal{E}_f} \dfrac{\|p_i - e_j\|}{u_i + v_j}$.

[1] https://youtu.be/H05SUfotwPc

Lemma 7. *At a given time instant $t \geq 0$, $i^* \subseteq \mathscr{F}(j^*)$, where $(i^*, j^*) = \underset{(i,j)\in\mathcal{P}\times\mathcal{E}_f}{\operatorname{argmin}} \dfrac{\|p_i - e_j\|}{u_i + v_j}$, and $\mathscr{F}$ is the final allocation function of A2 algorithm.*

Proof. From Definition 6 it is understood that t_s represents the minimum possible time taken to capture an agent in the evading team by an agent in the pursuing team, when the corresponding evader goes head-on with the pursuer. Therefore, if it can be shown that the Apollonius boundary around the evader j^* contains part of the Apollonius circle $\mathcal{A}_{i^*j^*}$ (i.e., $\mathcal{B}^t_{j^*} \cap \mathcal{A}_{i^*j^*} \neq \varnothing, \forall\, t \geq 0$), then the pursuer i^* will be assigned to the evader j^*, even when there is a tie. Since the Apollonius circle denotes the capture points for a non-maneuvering evader, the point $T_{i^*j^*}$, with the length of the line segment $E_{j^*}T_{i^*j^*} = v_{j^*}t_s$, is the closest capture point to evader j^* along the corresponding line of sight. Therefore $T_{i^*j^*} \in \mathcal{B}^t_{j^*}$, and hence, $\mathcal{B}^t_{j^*} \cap \mathcal{A}_{i^*j^*} \neq \varnothing, \forall\, t \geq 0$. □

Lemma 8. *Assuming the pursuers are assigned to the evaders using A2, at any given time $t \geq 0$, CST will converge to zero in finite time, and hence at least one evader will be captured in finite time.*

Proof. From Lemma 7, pursuer i^* (corresponding to the CST) is always assigned to evader j^*. Since all the pursuers are faster compared to the evaders, and since they follow a constant bearing strategy, $\mathrm{d}t_s/\mathrm{d}t \leq (v_{j^*} - u_{i^*})/(v_{j^*} + u_{i^*}) < 0$, for all $t \geq 0$ [167]. Furthermore, as the initial CST is finite, the CST converges to zero in finite time. Hence, capture of one evader is guaranteed in finite time. □

Theorem 4. *The A2 algorithm guarantees capture of all the evaders in finite time.*

Proof. The result immediately follows from Lemma 7 and Lemma 8. Note that the CST is updated (from zero) every time a capture occurs, and the captured evader is removed from the list of participating players. Also, the number of evaders are finite. □

CHAPTER 4
COVARIANCE STEERING FOR A CLASS OF LINEAR QUADRATIC STOCHASTIC DYNAMIC GAMES

In this chapter, we address the problem of steering a discrete-time linear dynamical system from an initial Gaussian distribution to a final distribution in a game-theoretic setting. One of the two players strives to minimize a quadratic payoff, while at the same time tries to meet a given mean and covariance constraint at the final time-step. The other player maximizes the same payoff, but it is assumed to be *indifferent* to the terminal constraint. At first, the unconstrained version of the game is examined, and the necessary conditions for the existence of a saddle point are obtained. It is then shown that obtaining a solution for the *one-sided* constrained dynamic game is not guaranteed, and subsequently the players' best responses are analyzed. The constrained covariance steering game is proposed to be solved numerically using the well-known Jacobi procedure. The problem of guiding a missile during the endgame is chosen to analyze the proposed approach. A test example corresponding to the case where the terminal distribution is not achieved is also included while discussing the necessary conditions to meet a given terminal constraint.

4.1 Mathematical Preliminaries

Consider the following discrete-time linear stochastic system

$$x_{k+1} = A_k x_k + B_k u_k + C_k v_k + D_k w_k, \tag{4.1}$$

where $k = 0,\ 1, \ldots, N-1$ is the time-step. At the k^{th} time-step, $x_k \in \mathbb{R}^n$ denotes the state, $u_k \in \mathbb{R}^m$ is the controller input, $v_k \in \mathbb{R}^\ell$ is the stopper input, and $w_k \in \mathbb{R}^r$ is a zero-mean

white Gaussian noise with unit covariance, i.e.

$$\mathbb{E}[w_k] = 0, \quad \mathbb{E}[w_{k_1} w_{k_2}^\top] = \begin{cases} I_r, & \text{if } k_1 = k_2, \\ 0, & \text{otherwise.} \end{cases} \tag{4.2}$$

In addition, it is assumed that

$$\mathbb{E}[x_{k_1} w_{k_2}^\top] = 0, \quad 0 \leq k_1 \leq k_2 \leq N. \tag{4.3}$$

The initial state x_0 is distributed according to $x_0 \sim \mathcal{N}(\mu_0, \Sigma_0)$, where $\mu_0 \in \mathbb{R}^n$ is the initial state mean, and $\Sigma_0 \in \mathbb{R}^{n \times n}$ is the initial state covariance, with $\Sigma_0 \succeq 0$. The payoff function is

$$J(u_0, \ldots, u_{N-1}, v_0, \ldots, v_{N-1}) = \mathbb{E}\left[\sum_{k=0}^{N-1} \left(x_k^\top Q_k x_k + u_k^\top R_k u_k - v_k^\top S_k v_k\right)\right]. \tag{4.4}$$

It is assumed that $Q_k \succeq 0$ for all $k = 0, \ldots, N$, and $R_k, S_k \succ 0$ for all $k = 0, \ldots, N-1$. The set of control inputs $\{u_0, \ldots, u_{N-1}\}$ is chosen by one player to minimize the payoff function (Equation 4.4), and the control inputs $\{v_0, \ldots, v_{N-1}\}$, are chosen by the adversary to maximize the payoff.

Using the notation introduced in [99], the system dynamics in Equation 4.1 can be alternatively expressed as

$$x_k = \bar{A}_k x_0 + \bar{B}_k U_k + \bar{C}_k V_k + \bar{D}_k W_k, \tag{4.5}$$

where $U_k = [u_0,\ u_1, \ldots, u_{k-1}]^\top$, $V_k = [v_0,\ v_1, \ldots, v_{k-1}]^\top$, $W_k = [w_0,\ w_1, \ldots, w_{k-1}]^\top$ are the augmented control and noise profiles. Furthermore, with the augmented state vector

$X = [x_1,\ x_2, \ldots, x_N]^\top$, the system dynamics (Equation 4.1) can be rewritten as

$$X = \mathcal{A}x_0 + \mathcal{B}U + \mathcal{C}V + \mathcal{D}W, \tag{4.6}$$

where $U = U_{N-1}$, $V = V_{N-1}$, and $W = W_{N-1}$. The definitions of the big matrices (A_k, $\mathcal{A}, \ldots$) can be found in [99]. Note that $\mathbb{E}[x_0 x_0^\top] = \Sigma_0 + \mu_0\mu_0^\top$, $\mathbb{E}[x_0 W^\top] = 0$, $\mathbb{E}[WW^\top] = I$. Consequently, the payoff function in Equation 4.4 can be expressed as

$$J(U, V) = \mathbb{E}[X^\top \bar{Q} X + U^\top \bar{R} U - V^\top \bar{S} V], \tag{4.7}$$

where $\bar{Q} = \text{blkdiag}(Q_0, \ldots, Q_{N-1}, 0) \in \mathbb{R}^{(N+1)n \times (N+1)n}$, $\bar{R} = \text{blkdiag}(R_0, R_1, \ldots, R_{N-1}) \in \mathbb{R}^{Nm \times Nm}$, and $\bar{S} = \text{blkdiag}(S_0, S_1, \ldots, S_{N-1}) \in \mathbb{R}^{N\ell \times N\ell}$. Also, since $Q_k \succeq 0$ for all $k = 0, \ldots, N$, and $R_k, S_k \succ 0$ for all $k = 0, \ldots, N-1$, it follows that $\bar{Q} \succeq 0$ and $\bar{R}, \bar{S} \succ 0$. The mean and the covariance of the initial state x_0 can be written in terms of X as

$$\mu_0 = E_0 \mathbb{E}[X], \tag{4.8a}$$

$$\Sigma_0 = E_0(\mathbb{E}[XX^\top] - \mathbb{E}[X]\mathbb{E}[X]^\top)E_0^\top, \tag{4.8b}$$

where $E_0 \triangleq [I_n, 0, \ldots, 0] \in \mathbb{R}^{n \times (N+1)n}$.

Definition 7. *The* upper game *is a scheme in which the stopper chooses V based on the information it has on the control U, and the* upper value *is defined by*

$$\mathcal{V}^+ = \inf_{U \in \mathbb{R}^{Nm}} \sup_{V \in \mathbb{R}^{N\ell}} J(U, V). \tag{4.9}$$

Similarly, the lower game *is a scheme in which the controller chooses U based on the*

information it has on the control V, and the lower value *is defined by*

$$\mathcal{V}^- = \sup_{V \in \mathbb{R}^{N\ell}} \inf_{U \in \mathbb{R}^{Nm}} J(U, V). \tag{4.10}$$

It is well known that, in general $\mathcal{V}^- \leq \mathcal{V}^+$. If the *Isaacs minimax condition* holds, then $\mathcal{V}^- = \mathcal{V}^+$, and the corresponding set of control actions (U^*, V^*) is called the equilibrium solution or saddle point [176]. The unconstrained Gaussian steering problems to be addressed in this paper can now be stated as follows.

Problem 1. *Find the saddle point (U^*, V^*) for the* unconstrained dynamic game *(UDG), described by the payoff function in Equation 4.7, the system in Equation 4.6, and the initial conditions in Equation 4.8.*

In this paper, as mentioned earlier, we propose to analyze the one-sided constrained dynamic game. To this end, let

$$E_N X = x_N \sim \mathcal{N}(\mu_N, \Sigma_N), \tag{4.11}$$

where $E_N \triangleq [0, \ldots, 0, I_n] \in \mathbb{R}^{n \times (N+1)n}$, be terminal state that the controller strives to achieve at the final time-step. Note that it is only the controller who is concerned about meeting the terminal condition in Equation 4.11, and hence Equation 4.11 is a one-sided constraint. It is assumed that the stopper is aware of the controller's terminal constraint however, it is indifferent to this constraint, and it is solely interested in maximizing the payoff (Equation 4.7). Furthermore, since the terminal constraint (Equation 4.11) is dependent on the control inputs of both players, and it is a one-sided constraint, the problem of interest can be categorized as a GCG [153].

Remark 2. *The terminal condition (Equation 4.11) can be used to enforce probabilistic capture in the case of a two-player pursuit-evasion game with $\mu_N = 0$, when Equation 4.1 represents the relative motion between the pursuer and the evader.*

We will now formally define the saddle point in the one-sided constrained dynamic game using the corresponding upper and lower values. For a given stopper action V, let $\mathcal{U}(V)$ denotes the set of controllers $U \in \mathbb{R}^{Nm}$ that drive the system to terminal Gaussian distribution in Equation 4.11, and let $\mathcal{R} \triangleq \bigcup_{V \in \mathbb{R}^{N\ell}} \mathcal{U}(V) \subseteq \mathbb{R}^{Nm}$.

Definition 8. *The* constrained upper value *is defined by*

$$\mathcal{V}_c^+ = \inf_{U \in \mathbb{R}^{Nm}} \sup_{V \in \mathbb{R}^{N\ell}} J(U, V), \tag{4.12}$$

and the constrained lower value *is defined by*

$$\mathcal{V}_c^- = \sup_{V \in \mathbb{R}^{N\ell}} \inf_{U \in \mathcal{R}} J(U, V). \tag{4.13}$$

The existence of the constrained upper and lower values requires that the controller meets the terminal constraint in Equation 4.11 in the corresponding upper and lower games. Given the system dynamics in Equation 4.6 and the initial conditions in Equation 4.8, note that for some V, there may not exist a controller such that the terminal condition (Equation 4.11) can be met i.e., $\mathcal{U}(V) = \emptyset$. Consequently, there may not exist a constrained upper (or lower) value for the constrained game. Finally, a saddle point in the constrained game can be defined as (U_c^*, V_c^*) for which the $\mathcal{V}_c^+$ and $\mathcal{V}_c^-$ exist, and are equal.

Problem 2. *Find the necessary conditions such that the controller can drive the system to the final state, while the stopper tries to maximize the payoff function (Equation 4.7), given the system dynamics (Equation 4.6) and the initial conditions (Equation 4.8). Furthermore, find the optimal control inputs for both players. Hereafter, this problem will be referred to as the* constrained dynamic game *(CDG).*

4.2 Separation of Mean and Covariance Control Problems

In Ref. [87], it was demonstrated that the mean and the covariance evolution of the system can be separated. Subsequently, by separating the cost, independent controllers that drive the mean and the covariance were derived. A similar approach is followed here by first observing the fact that

$$\mu_k \triangleq \mathbb{E}[x_k] = \bar{A}_k \mu_0 + \bar{B}_k \bar{U}_k + \bar{C}_k \bar{V}_k, \tag{4.14}$$

where $\bar{U}_k = \mathbb{E}[U_k]$ and $\bar{V}_k = \mathbb{E}[V_k]$. By defining

$$\tilde{x}_k \triangleq x_k - \mu_k, \quad \tilde{U}_k \triangleq U_k - \bar{U}_k, \quad \tilde{V}_k \triangleq V_k - \bar{V}_k, \tag{4.15}$$

and using Equation 4.5, the following equation holds for $\tilde{x}_k$.

$$\tilde{x}_k = \bar{A}_k \tilde{x}_0 + \bar{B}_k \tilde{U}_k + \bar{C}_k \tilde{V}_k + \bar{D}_k W_k. \tag{4.16}$$

Subsequently,

$$\begin{aligned} \Sigma_k &\triangleq \mathbb{E}[\tilde{x}_k \tilde{x}_k^\top] \\ &= \mathbb{E}\left[\left(\bar{A}_k \tilde{x}_0 + \bar{B}_k \tilde{U}_k + \bar{C}_k \tilde{V}_k + \bar{D}_k W_k \right) \left(\bar{A}_k \tilde{x}_0 + \bar{B}_k \tilde{U}_k + \bar{C}_k \tilde{V}_k + \bar{D}_k W_k \right)^\top \right]. \end{aligned} \tag{4.17}$$

It can be observed that the mean evolution in Equation 4.14 depends only on $\bar{U}_k$, $\bar{V}_k$, while the evolution of $\tilde{x}$ in Equation 4.16, and the covariance evolution in Equation 4.17 depend only on $\tilde{U}_k$, $\tilde{V}_k$, and the noise profile W_k. Consequently, from Equation 4.6 and Equation 4.14, it follows that

$$\bar{X} \triangleq \mathbb{E}[X] = \mathcal{A}\mu_0 + \mathcal{B}\bar{U} + \mathcal{C}\bar{V}, \tag{4.18}$$

and from Equation 4.16,

$$\tilde{X} \triangleq X - \mathbb{E}[X] = \mathcal{A}\tilde{x}_0 + \mathcal{B}\tilde{U} + \mathcal{C}\tilde{V} + \mathcal{D}W. \tag{4.19}$$

The objective function (Equation 4.7) can be further rewritten as

$$\begin{aligned} J(U,V) &= \mathbb{E}[X^\top \bar{Q} X + U^\top \bar{R} U - V^\top \bar{S} V] \\ &= \text{tr}(\bar{Q}\mathbb{E}[\tilde{X}\tilde{X}^\top]) + \bar{X}^\top \bar{Q}\bar{X} + \text{tr}(\bar{R}\mathbb{E}[\tilde{U}\tilde{U}^\top]) + \bar{U}^\top \bar{R}\bar{U} - \text{tr}(\bar{S}\mathbb{E}[\tilde{V}\tilde{V}^\top]) - \bar{V}^\top \bar{S}\bar{V} \\ &= J_\mu(\bar{U},\bar{V}) + J_\Sigma(\tilde{U},\tilde{V}), \end{aligned} \tag{4.20}$$

where

$$J_\mu(\bar{U},\bar{V}) = \bar{X}^\top \bar{Q}\bar{X} + \bar{U}^\top \bar{R}\bar{U} - \bar{V}^\top \bar{S}\bar{V}, \tag{4.21}$$

and

$$J_\Sigma(\tilde{U},\tilde{V}) = \text{tr}(\bar{Q}\mathbb{E}[\tilde{X}\tilde{X}^\top]) + \text{tr}(\bar{R}\mathbb{E}[\tilde{U}\tilde{U}^\top]) - \text{tr}(\bar{S}\mathbb{E}[\tilde{V}\tilde{V}^\top]). \tag{4.22}$$

Proposition 7. *For the UDG, the saddle point controls* (U^*, V^*) *that solve the problem (if they exist) are given by* $U^* = \bar{U}^* + \tilde{U}^*$ *and* $V^* = \bar{V}^* + \tilde{V}^*$*, where* $(\bar{U}^*, \bar{V}^*)$ *solves the* unconstrained mean steering game

$$\textit{(UMSG)} \begin{cases} \textit{Payoff function: } J_\mu(\bar{U},\bar{V}), \\ \textit{where } \bar{X} = \mathcal{A}\mu_0 + \mathcal{B}\bar{U} + \mathcal{C}\bar{V}, \end{cases} \tag{4.23}$$

and $(\tilde{U}^*, \tilde{V}^*)$ *solves the* unconstrained covariance steering game

$$\textit{(UCSG)} \begin{cases} \textit{Payoff function: } J_\Sigma(\tilde{U}, \tilde{V}), \\ \textit{where } \tilde{X} = \mathcal{A}\tilde{x}_0 + \mathcal{B}\tilde{U} + \mathcal{C}\tilde{V} + \mathcal{D}W. \end{cases} \tag{4.24}$$

Proof. From Equation 4.15 and Equation 4.20, it can observed that the mean payoff function $J_\mu(\bar{U}, \bar{V})$ in Equation 4.21 is driven by the mean control actions $\bar{U}$ and $\bar{V}$ independently while the covariance payoff $J_\Sigma(\tilde{U}, \tilde{V})$ in Equation 4.21 is driven by the covariance control actions $\tilde{U}$ and $\tilde{V}$. As a result, the UDG in terms of (U, V) is equivalent to two separate dynamic games in terms of $(\bar{U}, \bar{V})$ and $(\tilde{U}, \tilde{V})$ with payoff functions in Equation 4.21 and Equation 4.22, respectively, leading to the result. □

Proposition 8. *Proposition 7 applies to the CDG as well, with* $U_c^* = \bar{U}_c^* + \tilde{U}_c^*$, $V_c^* = \bar{V}_c^* + \tilde{V}_c^*$, *where* $(\bar{U}_c^*, \bar{V}_c^*)$ *solves the* constrained mean steering game *(CMSG)*

$$\begin{cases} \textit{Payoff function: } J_\mu(\bar{U}, \bar{V}), & \text{(4.25a)} \\ \textit{where } \bar{X} = \mathcal{A}\mu_0 + \mathcal{B}\bar{U} + \mathcal{C}\bar{V}, & \\ \textit{Controller constraint:} \mu_N = E_N\bar{X} = \bar{A}_N\mu_0 + \bar{B}_N\bar{U} + \bar{C}_N\bar{V}, & \text{(4.25b)} \end{cases}$$

and $(\tilde{U}_c^*, \tilde{V}_c^*)$ *solves the* constrained covariance steering game *(CCSG) with*

$$\begin{cases} \textit{Payoff function: } J_\Sigma(\tilde{U}, \tilde{V}), & \text{(4.26a)} \\ \textit{where } \tilde{X} = \mathcal{A}\tilde{x}_0 + \mathcal{B}\tilde{U} + \mathcal{C}\tilde{V} + \mathcal{D}W, & \\ \textit{Controller constraint: } \Sigma_N = E_N \left(\mathbb{E}[XX^\top] - \mathbb{E}[X]\mathbb{E}[X]^\top \right) E_N^\top, & \text{(4.26b)} \end{cases}$$

where the constraints (Equation 4.25b and Equation 4.26b), as stated earlier, are of concern only for the controller.

Proof. The proof is similar to the one given for Proposition 7. From Equation 4.15 and Equation 4.20, it can observed that the mean payoff $J_\mu(\bar{U}, \bar{V})$ in Equation 4.21 is driven by mean control actions $\bar{U}$ and $\bar{V}$, while the covariance payoff $J_\Sigma(\tilde{U}, \tilde{V})$ in Equation 4.21 is

driven by covariance control actions $\tilde{U}$ and $\tilde{V}$, independently. Furthermore, the mean and the covariance control actions address the constraints (Equation 4.25b and Equation 4.26b) too in an independent fashion. As a result, the CDG in terms of (U, V) is equivalent to two separate dynamic games in Equation 4.25 and Equation 4.26 in terms of $(\bar{U}, \bar{V})$ and $(\tilde{U}, \tilde{V})$, respectively, leading to the result. □

Note that non-existence of saddle point in either CMSG or CCSG or both, implies non-existence of saddle point in CDG. For the analysis of mean steering game in the following section, we introduce the set $\bar{\mathcal{R}}$. For a given stopper action $\bar{V}$ in CMSG, let $\bar{\mathcal{U}}(\bar{V})$ denotes the set of mean controllers $\bar{U} \in \mathbb{R}^{Nm}$ that satisfies the constraint in Equation 4.25b, and let $\bar{\mathcal{R}} \triangleq \bigcup_{\bar{V} \in \mathbb{R}^{N\ell}} \bar{\mathcal{U}}(\bar{V}) \subseteq \mathbb{R}^{Nm}$.

4.3 Mean Steering Game

The solution to the UMSG is given in the following proposition.

Proposition 9. *Assume that*

$$\bar{S} - \mathcal{C}^\top \bar{Q} \mathcal{C} \succ 0, \tag{4.27}$$

then the saddle point $(\bar{U}^*, \bar{V}^*)$ *that solves the UMSG (Equation 4.23) is given by*

$$\begin{bmatrix} \bar{U}^* \\ \bar{V}^* \end{bmatrix} = - \begin{bmatrix} \mathcal{B}^\top \bar{Q} \mathcal{B} + \bar{R} & \mathcal{B}^\top \bar{Q} \mathcal{C} \\ \mathcal{C}^\top \bar{Q} \mathcal{B} & \mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S} \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{B}^\top \bar{Q} \mathcal{A} \\ \mathcal{C}^\top \bar{Q} \mathcal{A} \end{bmatrix} \mu_0 \tag{4.28}$$

and this solution is unique.

Proof. The payoff function in Equation 4.21 can be expressed as

$$J_\mu(\bar{U}, \bar{V}) = (\mathcal{A}\mu_0 + \mathcal{B}\bar{U} + \mathcal{C}\bar{V})^\top \bar{Q} (\mathcal{A}\mu_0 + \mathcal{B}\bar{U} + \mathcal{C}\bar{V}) + \bar{U}^\top \bar{R} \bar{U} - \bar{V}^\top \bar{S} \bar{V}. \tag{4.29}$$

The first-order necessary conditions [177] for a saddle point yield

$$\nabla_{\bar{U}} J_\mu = (\mathcal{B}^\top \bar{Q} \mathcal{B} + \bar{R})\bar{U} + \mathcal{B}^\top \bar{Q} \mathcal{C} \bar{V} + \mathcal{B}^\top \bar{Q} \mathcal{A} \mu_0 = 0, \tag{4.30a}$$

$$\nabla_{\bar{V}} J_\mu = (\mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S})\bar{V} + \mathcal{C}^\top \bar{Q} \mathcal{B} \bar{U} + \mathcal{C}^\top \bar{Q} \mathcal{A} \mu_0 = 0. \tag{4.30b}$$

The above two equations can be expressed as

$$\begin{bmatrix} \mathcal{B}^\top \bar{Q} \mathcal{B} + \bar{R} & \mathcal{B}^\top \bar{Q} \mathcal{C} \\ \mathcal{C}^\top \bar{Q} \mathcal{B} & \mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S} \end{bmatrix} \begin{bmatrix} \bar{U}^* \\ \bar{V}^* \end{bmatrix} = - \begin{bmatrix} \mathcal{B}^\top \bar{Q} \mathcal{A} \\ \mathcal{C}^\top \bar{Q} \mathcal{A} \end{bmatrix} \mu_0, \tag{4.31}$$

Let

$$\mathcal{T}_m = \begin{bmatrix} \mathcal{B}^\top \bar{Q} \mathcal{B} + \bar{R} & \mathcal{B}^\top \bar{Q} \mathcal{C} \\ \mathcal{C}^\top \bar{Q} \mathcal{B} & \mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S} \end{bmatrix}, \tag{4.32}$$

and from Equation 4.27, $\mathcal{B}^\top \bar{Q} \mathcal{C} (\mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S})^{-1} \mathcal{C}^\top \bar{Q} \mathcal{B} \prec 0$. As a result, $\mathcal{B}^\top \bar{Q} \mathcal{B} + \bar{R} - \mathcal{B}^\top \bar{Q} \mathcal{C} (\mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S})^{-1} \mathcal{C}^\top \bar{Q} \mathcal{B} \succ 0$. Therefore, $\det(\mathcal{T}_m) = \det(\mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S}) \det(\mathcal{B}^\top \bar{Q} \mathcal{B} + \bar{R} - \mathcal{B}^\top \bar{Q} \mathcal{C} (\mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S})^{-1} \mathcal{C}^\top \bar{Q} \mathcal{B}) \neq 0$, and $\mathcal{T}_m$ is invertible. Equation 4.28 then follows immediately from Equation 4.31. From Equation 4.27, the second order derivatives yield

$$\nabla_{\bar{U}\bar{U}} J_\mu = \mathcal{B}^\top \bar{Q} \mathcal{B} + \bar{R} \succ 0, \tag{4.33a}$$

$$\nabla_{\bar{V}\bar{V}} J_\mu = \mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S} \prec 0. \tag{4.33b}$$

Therefore, the payoff function is convex in $\bar{U}$, and concave in $\bar{V}$. Hence $(\bar{U}^*, \bar{V}^*)$ is the only saddle point that solves the given dynamic game [177]. □

Next, we analyze the CMSG. As this is a constrained zero-sum game, we obtain the following inequality. A similar result can be found in Ref. [153] (Theorem III.1).

Lemma 9. *Assuming that the UMSG (Equation 4.23) has a saddle point equilibrium (Propo-*

sition 9), the CMSG (Equation 4.25) satisfies

$$\inf_{\bar{U}\in\mathbb{R}^{Nm}} \sup_{\bar{V}\in\mathbb{R}^{N\ell}} J_\mu(\bar{U},\bar{V}) \leq \sup_{\bar{V}\in\mathbb{R}^{N\ell}} \inf_{\bar{U}\in\bar{\mathcal{R}}} J_\mu(\bar{U},\bar{V}). \tag{4.34}$$

Proof. Given that the UMSG has a saddle point equilibrium, it follows that

$$\inf_{\bar{U}\in\mathbb{R}^{Nm}} \sup_{\bar{V}\in\mathbb{R}^{N\ell}} J_\mu(\bar{U},\bar{V}) = \sup_{\bar{V}\in\mathbb{R}^{N\ell}} \inf_{\bar{U}\in\mathbb{R}^{Nm}} J_\mu(\bar{U},\bar{V}). \tag{4.35}$$

Since $\bar{\mathcal{R}} \subseteq \mathbb{R}^{Nm}$,

$$\inf_{\bar{U}\in\mathbb{R}^{Nm}} J_\mu(\bar{U},\bar{V}) \leq \inf_{\bar{U}\in\bar{\mathcal{R}}} J_\mu(\bar{U},\bar{V}). \tag{4.36}$$

Hence,

$$\sup_{\bar{V}\in\mathbb{R}^{N\ell}} \inf_{\bar{U}\in\mathbb{R}^{Nm}} J_\mu(\bar{U},\bar{V}) \leq \sup_{\bar{V}\in\mathbb{R}^{N\ell}} \inf_{\bar{U}\in\bar{\mathcal{R}}} J_\mu(\bar{U},\bar{V}), \tag{4.37}$$

and from Equation 4.35, the result follows. □

As a result, a pure-strategy equilibrium might not exist for the CMSG, and only players' best responses can be obtained [153]. To this end, the constrained upper and lower games for the CMSG problem can be examined. As stated in Definition 8, in the constrained lower game, the stopper has to choose its input first, while the controller has the advantage of obtaining the stopper input, and then choosing his best response accordingly.

Lemma 10. *Assuming that the discrete-time linear dynamical system in Equation 4.1 is controllable for $C_k = 0$ and $D_k = 0$ (i.e., rank$[\bar{B}_N] = n$), the controller's feasible set (the set of controllers for which the constraint in Equation 4.25b is met given the stopper input) is non-empty for any $\bar{V} \in \mathbb{R}^{N\ell}$.*

Proof. For a given $\bar{V} \in \mathbb{R}^{N\ell}$, the mean constraint (Equation 4.25b) can be rewritten as

$$\bar{B}_N \bar{U} = \mu_N - \bar{A}_N \mu_0 - \bar{C}_N \bar{V}. \tag{4.38}$$

Since μ_0 and μ_N are known, and since rank$[\bar{B}_N] = n$, there exists a solution for $\bar{U}$ for every $\bar{V} \in \mathbb{R}^{N\ell}$. Hence, the controller's feasible set is non-empty. □

From the above lemma, it is obvious that the controller can meet the mean constraint (Equation 4.25b), if the condition rank$[\bar{B}_N] = n$ is satisfied. In the upper game, the controller input is obtained first and the stopper best responds accordingly. The terminal condition in Equation 4.25b depends on the stopper input. Note that it is assumed that the stopper is indifferent to this constraint, and in this regard, the sufficient condition for which the controller's terminal constraint is met is derived in Lemma 11 below.

From Equation 4.30a and Equation 4.30b, the players' best responses as a function of their opponent response can be obtained as

$$\bar{U} = -(\mathcal{B}^\top \bar{Q} \mathcal{B} + \bar{R})^{-1} (\mathcal{B}^\top \bar{Q} \mathcal{C} \bar{V} + \mathcal{B}^\top \bar{Q} \mathcal{A} \mu_0), \tag{4.39a}$$

$$\bar{V} = -(\mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S})^{-1} (\mathcal{C}^\top \bar{Q} \mathcal{B} \bar{U} + \mathcal{C}^\top \bar{Q} \mathcal{A} \mu_0). \tag{4.39b}$$

In the upper game, where the controller plays first, the stopper input as a function of $\bar{U}$ is given by Equation 4.39b. Given the stopper input (as per Equation 4.39b), from the constraint in Equation 4.25b, it follows that

$$\begin{aligned} \mu_N &= \bar{A}_N \mu_0 + \bar{B}_N \bar{U} + \bar{C}_N \big(- (\mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S})^{-1} (\mathcal{C}^\top \bar{Q} \mathcal{B} \bar{U} + \mathcal{C}^\top \bar{Q} \mathcal{A} \mu_0) \big) \\ &= \big(\bar{A}_N - \bar{C}_N (\mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S})^{-1} \mathcal{C}^\top \bar{Q} \mathcal{A} \big) \mu_0 + \big(\bar{B}_N - \bar{C}_N (\mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S})^{-1} \mathcal{C}^\top \bar{Q} \mathcal{B} \big) \bar{U}. \end{aligned} \tag{4.40}$$

For the sake of brevity, let $\mathcal{G} = \bar{B}_N - \bar{C}_N (\mathcal{C}^\top \bar{Q} \mathcal{C} - \bar{S})^{-1} \mathcal{C}^\top \bar{Q} \mathcal{B}$.

Lemma 11. *Given the CMSG (Equation 4.25), in the associated upper game, the constraint in Equation 4.25b is satisfied if and only if*

$$\textit{rank}\left[\mathcal{G} \quad \mu_N - \left(\bar{A}_N - \bar{C}_N(\mathcal{C}^\top \bar{Q}\mathcal{C} - \bar{S})^{-1}\mathcal{C}^\top \bar{Q}\mathcal{A}\right)\mu_0\right] = \textit{rank}\left[\mathcal{G}\right]. \tag{4.41}$$

Proof. The condition in Equation 4.41 suggests that the system of linear equations, obtained from Equation 4.40,

$$\mathcal{G}\bar{U} = \mu_N - \left(\bar{A}_N - \bar{C}_N(\mathcal{C}^\top \bar{Q}\mathcal{C} - \bar{S})^{-1}\mathcal{C}^\top \bar{Q}\mathcal{A}\right)\mu_0, \tag{4.42}$$

has a solution for $\bar{U}$. Therefore, there always exists a constrained upper value for the CMSG, and the controller can drive the state to a given μ_N at the N^{th} time-step. □

Note that the matrix $\mathcal{G}$ can be treated as a *relative controllability matrix*, similar to the one introduced in Ref. [178] for continuous systems. The optimal control sequences $\bar{U}_*$ and $\bar{V}_*$ that solve the upper game can be found as follows. From Equation 4.39b, the upper game can be expressed in terms of the following minimization problem.

$$\begin{cases} \min\limits_{\bar{U}\in\mathbb{R}^{Nm}} \bar{X}^\top \bar{Q}\bar{X} + \bar{U}^\top \bar{R}\bar{U} - \bar{V}^\top \bar{S}\bar{V}, \\ \text{subject to } \mu_N = \bar{A}_N\mu_0 + \bar{B}_N\bar{U} + \bar{C}_N\bar{V}, \end{cases} \tag{4.43}$$

where $\bar{X} = \mathcal{A}\mu_0 + \mathcal{B}\bar{U} + \mathcal{C}\bar{V}$, and $\bar{V} = -(\mathcal{C}^\top \bar{Q}\mathcal{C} - \bar{S})^{-1}(\mathcal{C}^\top \bar{Q}\mathcal{B}\bar{U} + \mathcal{C}^\top \bar{Q}\mathcal{A}\mu_0)$.

Proposition 10. *Under the assumption*

$$\textit{rank}\ \mathcal{G} = n, \tag{4.44}$$

the optimal control sequence $\bar{U}_*$ that solves the minimization problem in Equation 4.43 is

given by

$$\bar{U}_* = \mathcal{R}^{-1}\left(\mathcal{M} + \mathcal{G}^\top \lambda/2\right), \tag{4.45}$$

where

$$\mathcal{R} = \bar{R} + \mathcal{B}^\top \bar{Q}\mathcal{B} - \mathcal{B}^\top \bar{Q}\mathcal{C}(\mathcal{C}^\top \bar{Q}\mathcal{C} - \bar{S})^{-1}\mathcal{C}^\top \bar{Q}\mathcal{B}, \tag{4.46a}$$

$$\mathcal{M} = \left(\mathcal{B}^\top \bar{Q}\mathcal{C}(\mathcal{C}^\top \bar{Q}\mathcal{C} - \bar{S})^{-1}\mathcal{C}^\top - \mathcal{B}^\top\right)\bar{Q}\mathcal{A}\mu_0, \tag{4.46b}$$

$$\lambda = 2\left(\mathcal{G}\mathcal{R}^{-1}\mathcal{G}^\top\right)^{-1}\left(\mu_N - \bar{A}_N\mu_0 + \bar{C}_N(\mathcal{C}^\top \bar{Q}\mathcal{C} - \bar{S})^{-1}\mathcal{C}^\top \bar{Q}\mathcal{A}\mu_0 - \mathcal{G}\mathcal{R}^{-1}\mathcal{M}\right). \tag{4.46c}$$

Proof. The Lagrangian for the constrained minimization problem (Equation 4.43) can be written as

$$\begin{aligned}\mathcal{L}(\bar{U}, \lambda) &= \bar{X}^\top \bar{Q}\bar{X} + \bar{U}^\top \bar{R}\bar{U} - \bar{V}^\top \bar{S}\bar{V} + \lambda^\top(\mu_N - \bar{A}_N\mu_0 - \bar{B}_N\bar{U} - \bar{C}_N\bar{V}) \\ &= (\mathcal{A}\mu_0 + \mathcal{B}\bar{U} + \mathcal{C}\bar{V})^\top \bar{Q}(\mathcal{A}\mu_0 + \mathcal{B}\bar{U} + \mathcal{C}\bar{V}) + \bar{U}^\top \bar{R}\bar{U} - \bar{V}^\top \bar{S}\bar{V} \\ &\quad + \lambda^\top(\mu_N - \bar{A}_N\mu_0 - \bar{B}_N\bar{U} - \bar{C}_N\bar{V}),\end{aligned} \tag{4.47}$$

where $\lambda \in \mathbb{R}^n$. The first-order optimality condition yields

$$\begin{aligned}\nabla_{\bar{U}}\mathcal{L} = 2(\mathcal{A}\mu_0 + \mathcal{B}\bar{U} + \mathcal{C}\bar{V})^\top \bar{Q}\left(\mathcal{B} + \mathcal{C}\frac{\partial \bar{V}}{\partial \bar{U}}\right) + 2\bar{U}^\top \bar{R} - 2\bar{V}^\top \bar{S}\frac{\partial \bar{V}}{\partial \bar{U}} \\ + \lambda^\top\left(-\bar{B}_N - \bar{C}_N\frac{\partial \bar{V}}{\partial \bar{U}}\right) = 0,\end{aligned} \tag{4.48}$$

and Equation 4.45 follows from the fact that $\dfrac{\partial \bar{V}}{\partial \bar{U}} = -(\mathcal{C}^\top \bar{Q}\mathcal{C} - \bar{S})^{-1}\mathcal{C}^\top \bar{Q}\mathcal{B}$, obtained using

(Equation 4.39b), and from the second-order optimality condition

$$\begin{aligned}\frac{\nabla_{\bar{U}\bar{U}}\mathcal{L}}{2} &= \left(\mathcal{B}+\mathcal{C}\frac{\partial\bar{V}}{\partial\bar{U}}\right)^{\top}\bar{Q}\left(\mathcal{B}+\mathcal{C}\frac{\partial\bar{V}}{\partial\bar{U}}\right)+\bar{R}-\frac{\partial\bar{V}}{\partial\bar{U}}^{\top}\bar{S}\frac{\partial\bar{V}}{\partial\bar{U}} \\ &= \left(\bar{R}+\mathcal{B}^{\top}\bar{Q}\mathcal{B}-\mathcal{B}^{\top}\bar{Q}\mathcal{C}(\mathcal{C}^{\top}\bar{Q}\mathcal{C}-\bar{S})^{-1}\mathcal{C}^{\top}\bar{Q}\mathcal{B}\right) \\ &= \mathcal{R}\succ 0 \end{aligned} \tag{4.49}$$

The Lagrange multiplier λ can be found by substituting Equation 4.45 along with Equation 4.46a and Equation 4.46b into the terminal constraint, obtaining

$$\left(\mathcal{G}\mathcal{R}^{-1}\mathcal{G}^{\top}\right)\lambda = 2\left(\mu_N-\bar{A}_N\mu_0+\bar{C}_N(\mathcal{C}^{\top}\bar{Q}\mathcal{C}-\bar{S})^{-1}\mathcal{C}^{\top}\bar{Q}\mathcal{A}\mu_0-\mathcal{G}\mathcal{R}^{-1}\mathcal{M}\right) \tag{4.50}$$

Note that since $\mathcal{R}$ is invertible and $\mathcal{G}$ has full row rank, $\mathcal{G}\mathcal{R}^{-1}\mathcal{G}^{\top}$ is invertible. □

4.4 Covariance Steering Game

The methodology to solve the UCSG and the CCSG is presented in this section. Assuming a linear feedback control structure for steering the covariance, we express $\tilde{U}$ and $\tilde{V}$ as

$$\tilde{u}_k = K_k y_k, \quad \tilde{v}_k = L_k y_k, \tag{4.51}$$

where $K_k \in \mathbb{R}^{m\times n}$, $L_k \in \mathbb{R}^{\ell\times n}$,

$$y_{k+1} = A_k y_k + D_k w_k, \tag{4.52a}$$

$$y_0 = x_0 - \mu_0, \tag{4.52b}$$

and $y_k \in \mathbb{R}^n$. Note that $\mathbb{E}[y_0] = 0$ and $\mathbb{E}[y_0 y_0^{\top}] = \Sigma_0$. Further, it can be obtained that

$$Y = \mathcal{A}y_0 + \mathcal{D}W, \tag{4.53}$$

where $Y = [y_0^\top, \ldots, y_N^\top]^\top \in \mathbb{R}^{(N+1)n}$, using the matrices introduced in section 4.1. Therefore, $\tilde{X}$ in Equation 4.19 can be rewritten as

$$\tilde{X} = (I + \mathcal{B}K + \mathcal{C}L)(\mathcal{A}y_0 + \mathcal{D}W). \tag{4.54}$$

where,

$$K = \begin{bmatrix} K_0 & 0 & \ldots & 0 & 0 \\ 0 & K_1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & K_{N-1} & 0 \end{bmatrix}, \tag{4.55a}$$

$$L = \begin{bmatrix} L_0 & 0 & \ldots & 0 & 0 \\ 0 & L_1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & L_{N-1} & 0 \end{bmatrix}, \tag{4.55b}$$

are the controller and the stopper gain matrices, respectively. Here $K \in \mathbb{K}$ and $L \in \mathbb{L}$, where $\mathbb{K}$ is the set of $Nm \times (N+1)n$ matrices that have the structure shown in Equation 4.55a, and similarly, $\mathbb{L}$ is the set of $N\ell \times (N+1)n$ matrices that have the structure shown in Equation 4.55b. From Equation 4.51, Equation 4.53, and Equation 4.54, we have $\mathbb{E}[\tilde{X}\tilde{X}^\top] = (I + \mathcal{B}K + \mathcal{C}L)\Sigma_s(I + \mathcal{B}K + \mathcal{C}L)^\top$, $\mathbb{E}[\tilde{U}\tilde{U}^\top] = K\Sigma_s K^\top$, $\mathbb{E}[\tilde{V}\tilde{V}^\top] = L\Sigma_s L^\top$, where $\Sigma_s = \mathcal{A}\Sigma_0\mathcal{A}^\top + \mathcal{D}\mathcal{D}^\top$. Therefore, the cost function Equation 4.22 can be converted to the following quadratic form in terms of K and L:

$$J_\Sigma(K, L) = \text{tr}(((I + \mathcal{B}K + \mathcal{C}L)^\top \bar{Q}(I + \mathcal{B}K + \mathcal{C}L) + K^\top \bar{R}K - L^\top \bar{S}L)\Sigma_s), \tag{4.56}$$

and the terminal constraint (Equation 4.26b) can be rewritten as

$$\Sigma_N = E_N(I + \mathcal{B}K + \mathcal{C}L)\Sigma_s(I + \mathcal{B}K + \mathcal{C}L)^\top E_N^\top. \tag{4.57}$$

For the sake of analysis, we introduce the set $\tilde{\mathcal{R}}$. Given stopper gain L in CCSG, let $\mathcal{K}(L)$ denotes the set of gains $K \in \mathbb{K}$ for which the controller satisfies the constraint in Equation 4.57, and let $\tilde{\mathcal{R}} \triangleq \bigcup_{L\in\mathbb{L}} \mathcal{K}(L) \subseteq \mathbb{K}$.

We first analyze the UCSG. Since the gain matrices K and L have constraints on their structure with zeros, as shown in Equation 4.55, with a slight abuse of notation, the Lagrangian can be written as

$$\begin{aligned}\mathcal{L}(K, L, \Theta, \Xi) = \text{tr}(((I + \mathcal{B}K + \mathcal{C}L)^\top \bar{Q}(I + \mathcal{B}K + \mathcal{C}L) + K^\top \bar{R}K - L^\top \bar{S}L)\Sigma_s)/2 \\ + \sum_{i=1}^{Nm} \sum_{j\in\mathscr{J}_k(i)} \theta_{ij} e_i^\top K e_j + \sum_{i=1}^{N\ell} \sum_{j\in\mathscr{J}_l(i)} \xi_{ij} e_i^\top L e_j,\end{aligned} \tag{4.58}$$

where the functions $\mathscr{J}_k(.)$ and $\mathscr{J}_l(.)$ map each row number to the set of columns in which the gains K and L, respectively, have zero elements. The matrices $\Theta \in \mathbb{R}^{Nm\times(N+1)n}$ and $\Xi \in \mathbb{R}^{N\ell\times(N+1)n}$ are Lagrange multipliers of sizes equal to K and L, respectively. Note that the blocks in Θ and Ξ (corresponding to K_k and L_k) are zeros, and θ_{ij} and ξ_{ij} are the non-zero elements of these matrices. The first-order necessary conditions for the existence of a saddle point can be obtained by taking derivatives of the Lagrangian in Equation 4.58 with respect to K and L as

$$\nabla_K \mathcal{L} = \left[\mathcal{B}^\top \bar{Q} + \bar{R}K + \mathcal{B}^\top \bar{Q}\mathcal{B}K + \mathcal{B}^\top \bar{Q}\mathcal{C}L\right]\Sigma_s + \Theta = 0, \tag{4.59a}$$

$$\nabla_L \mathcal{L} = \left[\mathcal{C}^\top \bar{Q} - \bar{S}L + \mathcal{C}^\top \bar{Q}\mathcal{B}K + \mathcal{C}^\top \bar{Q}\mathcal{C}L\right]\Sigma_s + \Xi = 0. \tag{4.59b}$$

The candidate solutions for the UCSG can be obtained by solving the linear system of equations given in Equation 4.59. Since the gradients are linear, the second-order sufficient

conditions, using the bordered Hessians, can be invoked to find the saddle points among the candidate solutions numerically [179].

Next, we analyze the CCSG. A result similar to the one proposed for the CMSG (Lemma 9) follows for the CCSG and is given below. The proof is omitted as it is similar to the one given for Lemma 9.

Lemma 12. *Assuming that the UCSG with payoff function in Equation 4.56 has a saddle point equilibrium, then the CCSG (Equation 4.26), with the terminal constraint Equation 4.57 imposed only for the controller, satisfies*

$$\inf_{K\in\mathbb{K}} \sup_{L\in\mathbb{L}} J_\Sigma(K,L) \leq \sup_{L\in\mathbb{L}} \inf_{K\in\tilde{\mathcal{R}}} J_\Sigma(K,L). \tag{4.60}$$

Similarly, in the CCSG, a pure-strategy equilibrium need not exist. To this end, consider a simple Jacobi procedure given in Algorithm 2 to arrive at an equilibrium solution, assuming one exists. For Algorithm 2 to converge to an equilibrium solution for any K_0, L_0, the solution has to be a stable one [180]. The conditions for the existence of a stable equilibrium for the case where the cost is convex in K and concave in L can be found in Ref. [180].

Algorithm 2 Jacobi procedure to obtain saddle points

procedure JACOBI(K_0,L_0)
  for i = 0,1,2,... **do**
    $L_{i+1} := \arg\max_{L\in\mathbb{L}} J_\Sigma(K_i, L)$
    $K_{i+1} := \arg\min_{K\in\mathcal{K}(L_i)} J_\Sigma(K, L_i)$
  end for
  return K_{i+1}, L_{i+1}
end procedure

Subsequently, under the assumptions that $\Sigma_s \otimes (\mathcal{B}^\top \bar{Q}\mathcal{B} + \bar{R}) \succ 0$ (convex in K) and $\Sigma_s \otimes (\mathcal{C}^\top \bar{Q}\mathcal{C} - \bar{S}) \prec 0$ (concave in L), we can formulate the successive minimization and maximization problems as convex programming problems by relaxing the equality

constraint in Equation 4.57 to an inequality constraint,

$$\Sigma_N \succeq E_N(I + \mathcal{B}K + \mathcal{C}L)\Sigma_s(I + \mathcal{B}K + \mathcal{C}L)^\top E_N^\top. \tag{4.61}$$

Lemma 13. *Assuming $\Sigma_N \succ 0$, the inequality constraint in Equation 4.61 can be expressed as*

$$\|\Sigma_N^{-1/2} E_N(I + \mathcal{B}K + \mathcal{C}L)\Sigma_s^{1/2}\|_2 - 1 \leq 0. \tag{4.62}$$

Proof. Since assumption $\Sigma_N \succ 0$, Equation 4.61 can be rewritten as

$$I - \Sigma_N^{-1/2} E_N(I + \mathcal{B}K + \mathcal{C}L)\Sigma_s(I + \mathcal{B}K + \mathcal{C}L)^\top E_N^\top \Sigma_N^{-1/2} \succeq 0.$$

As it is symmetric, the matrix $\Sigma_N^{-1/2} E_N(I + \mathcal{B}K + \mathcal{C}L)\Sigma_s(I + \mathcal{B}K + \mathcal{C}L)^\top E_N^\top \Sigma_N^{-1/2}$ is diagonalizable via an orthogonal matrix $T \in \mathbb{R}^{n\times n}$ as

$$T(I_n - \operatorname{diag}(\lambda_1, \ldots, \lambda_n))S^\top \succeq 0, \tag{4.63}$$

where $\lambda_1, \ldots, \lambda_n$ are its eigenvalues. From Equation 4.63, we have

$$1 - \lambda_{max}\big(\Sigma_N^{-1/2} E_N(I + \mathcal{B}K + \mathcal{C}L)\Sigma_s(I + \mathcal{B}K + \mathcal{C}L)^\top E_N^\top \Sigma_N^{-1/2}\big) \geq 0. \tag{4.64}$$

$$\implies 1 - \|\Sigma_N^{-1/2} E_N(I + \mathcal{B}K + \mathcal{C}L)\Sigma_s^{1/2}\|_2 \geq 0. \tag{4.65}$$

Hence proved. □

4.5 Numerical Simulations

As mentioned earlier, in the lower game of the mean steering case, the controller has an advantage to drive the distribution to a given terminal Gaussian, assuming the system is

controllable. A more challenging case is that of the upper game, where the controller has to ensure that the terminal constraint (Equation 4.25b) is met while choosing its input first. In this section, we first present test examples for the upper game of the CMSG with linear time-invariant systems, and then analyze the missile end-game guidance problem. For the covariance steering part, YALMIP [181] in conjunction with MOSEK [182] was used to solve the successive convex optimization problems in the Jacobi procedure. The convergence criterion for the iterative method is $\epsilon_k, \epsilon_\ell \leq \epsilon$, where $\epsilon_k = \|K_{i+1} - K_i\|$ and $\epsilon_\ell = \|L_{i+1} - L_i\|$.

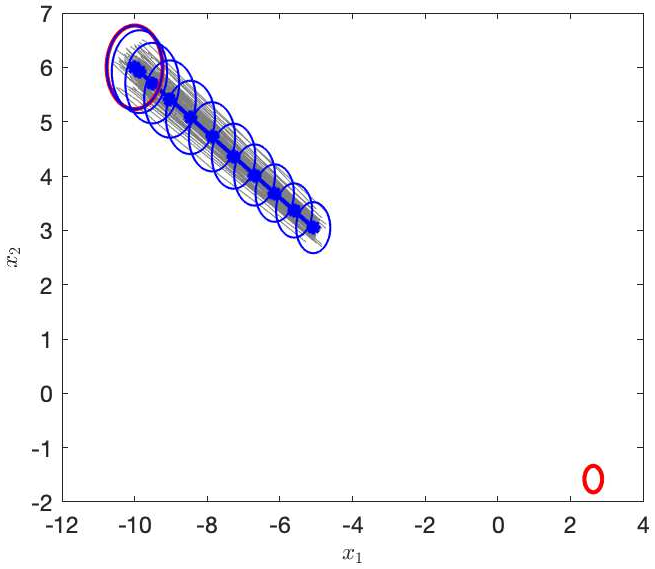


Figure 4.1: Unconstrained mean and covariance steering

4.5.1 Test Example

Consider the linear system

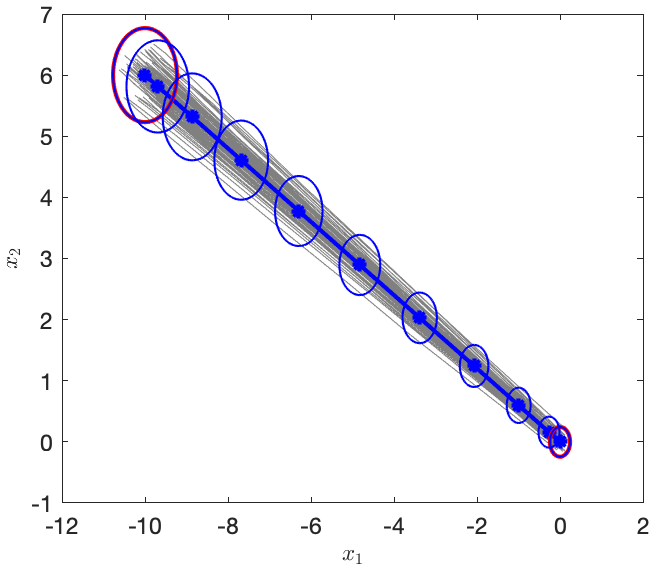
$$z_{k+1} = Az_k + Bu_k + Cv_k + Dw_k \tag{4.66}$$

where $z_k = [x_1, x_2, x_3, x_4]^\top \in \mathbb{R}^4$, $u_k, v_k \in \mathbb{R}^2$, $w_k \in \mathbb{R}^4$,

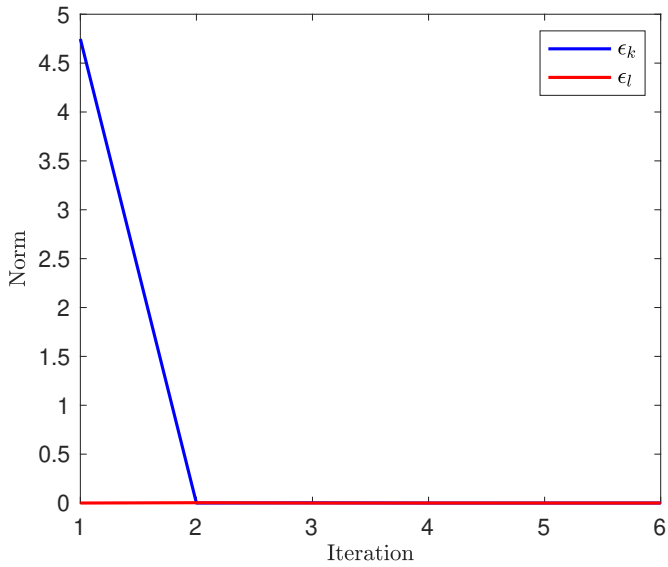
$$A = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} \Delta t^2 & 0 \\ 0 & \Delta t^2 \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix}, \tag{4.67}$$

$C = -B$, and $D = 0.01 I_4$. Note that x_1, x_2 can be understood as relative coordinates, and x_3, x_4 are the relative velocities along the x_1 and x_2 axes, respectively, with $\Delta t = 0.2$ being the time-step size. Finally, u_k and v_k are the accelerations of the pursuer (controller) and the evader (stopper), respectively.

The initial condition is chosen to be $\mu_0 = [-10, 6, 0, 0]^\top$, $\Sigma_0 =$ blkdiag(0.05, 0.05, 0.01, 0.01), and the terminal constraint is $\mu_N = [0, 0, 0, 0]^\top$, $\Sigma_N =$ blkdiag(0.005, 0.005, 0.001, 0.001). The time horizon is fixed at $N = 10$, and the cost matrices are $Q_k = I_4$, and $R_k = I_2$ $S_k = 100 I_2$, for all $k \geq 1$. The instance is first analyzed without the terminal constraint, and the solution to the unconstrained game is obtained. The UCSG is solved using the Jacobi procedure illustrated in Ref. [180]. Figure 4.1 presents the solution to the unconstrained game for the given initial condition. The red ellipses in Figure 4.1 denote the 3σ error of the initial and the desired terminal state distributions of x_1 and x_2 coordinates. The blue solid line illustrates the mean trajectory, and the blue ellipses illustrate the covariance evolution over the time horizon. The gray lines are the trajectories simulated for 100 different initial conditions that are sampled from $\mathcal{N}(\mu_0, \Sigma_0)$. It can be observed that the mean and the covariance trajectories do not meet the controller's terminal conditions. For the CDG, the relative controllability matrix is found to have full row rank, and therefore the mean can be driven to the specified terminal value. Also, the covariance steering problem is feasible with $\epsilon = 10^{-5}$, and the result is illustrated in Figure 4.2. The convergence of the Jacobi procedure (Algorithm 2) can be observed in Figure 4.2(b). From Figure 4.2(a), it can be observed that the covariance constraint is satisfied.



(a) Sampled Trajectories



(b) Convergence of the Jacobi procedure

Figure 4.2: Constrained mean and covariance steering: A case where the covariance condition is met by the controller

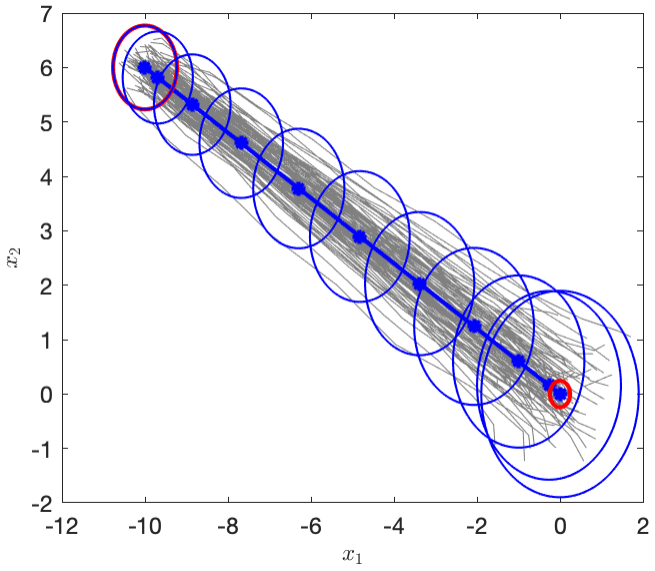


Figure 4.3: A case where the terminal covariance constraint is not met.

Figure 4.3 illustrates the case where $D = 0.1I_4$, while the rest of the values are kept unchanged. Since changing the matrix D does not change the behavior of the mean, in this case, the mean converges to the specified terminal value. However, the covariance constraint cannot be achieved in this case and from Figure 4.3, it can be observed that the covariance ellipse grows with time. The result in Figure 4.3 is for the set of optimal gains (K^*, L^*), obtained by minimizing the cost in Equation 4.22 subject to the constraint in Equation 4.59b, since the constraint in Equation 4.62 cannot be met.

4.5.2 Missile Endgame Guidance

Figure 4.4 presents a schematic of the interception geometry during a missile engagement scenario. During the endgame, the relative dynamics can be linearized along the initial line of sight while assuming a constant closing speed $V_c = V_p + V_e$, where V_p and V_e are the constant speeds of the missile and of the target, respectively. Trajectory linearization is well established for ballistic missile defense where the endgame is over a short duration, and begins with near "head-on" initial conditions. Now, consider the linearized dynamics

of a missile during the end-game in continuous time [183],

$$\dot{z} = A_c z + B_c u + C_c v, \tag{4.68}$$

where $z = [y, \dot{y}, a_e, a_p]^\top$, y is the target's relative distance to the missile normal to the reference line (initial line-of-sight), $\dot{y}$ is the relative speed, a_e and a_p are the lateral forces acting on the target and on the missile, respectively. In Equation 4.68, u and v are the commanded lateral accelerations of the missile and of the target, respectively. The matrices in Equation 4.68 are given by

$$A_c = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1/\tau_e & 0 \\ 0 & 0 & 0 & -1/\tau_p \end{bmatrix}, \quad B_c = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1/\tau_p \end{bmatrix}, \quad C_c = \begin{bmatrix} 0 \\ 0 \\ 1/\tau_e \\ 0 \end{bmatrix}, \tag{4.69a}$$

where τ_p, and τ_e are model parameters [183]. The corresponding discrete matrices can be obtained from a given time-step size Δt, and the system evolution is expressed using (Equation 4.66) by assuming process noise in the system entering through the control channels. The system matrices for this example are given by $A = \exp(A_c \Delta t)$, $B = \int_0^{\Delta t} \exp(A_c \tau) B_c \mathrm{d}\tau$, $C = \int_0^{\Delta t} \exp(A_c \tau) C_c \mathrm{d}\tau$,

$$D = \alpha \int_0^{\Delta t} \exp(A_c \tau) \mathrm{d}\tau \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{4.70}$$

where α is a constant.

For simulation purposes, the following parameters are chosen: $\Delta t = 0.1$, the initial separation along the line of sight $x_0 = 3500$, $\tau_e = 0.02$, $\tau_p = 0.01$. We choose $V_p = 3000$,

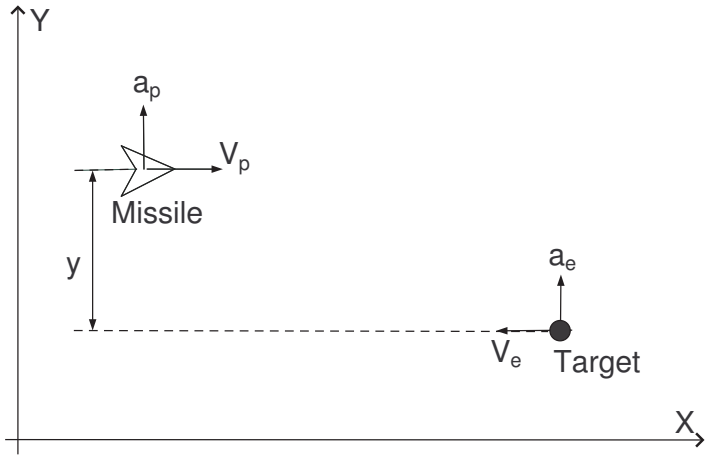


Figure 4.4: Planar end-game scenario

$V_e = 2000$, and therefore $t_f = x_0/(V_p + V_e) = 0.7$, $N = t_f/\Delta t = 7$. The initial conditions are $\mu_0 = [0, 350, 0, 0]^\top$, $\Sigma_0 = \text{blkdiag}(0.2, 0.2, 0.2, 0.2)$, and the terminal constraint is $\mu_N = [0, 0, 0.1, 0.1]^\top$, $\Sigma_N = \text{blkdiag}(0.1, 10, 1, 1)$. Furthermore, $Q_k = 10^{-6} I_4$, $R_k = 10^2$, $S_k = 3 \times 10^8$, for all $k \geq 1$. For this example, the relative controllability matrix is found to have full row rank, and the covariance steering problem is feasible. The result is illustrated in Figure 4.5, which shows the relative distance y versus the time-step. The errorbars (in red) indicate the 3σ error in y at the initial and the final time-steps. From Figure 4.5, it can be observed that the mean and covariance constraints are satisfied to successfully intercept the target during the end-game.

Figure 4.6 shows one of the many missile trajectories (generated for Figure 4.5) relative to the target (red cross) in the $x-y$ plane. Note that the motion along the y-axis is negligible compared to the missile's motion along the x-axis and consequently, a skewed aspect ratio has to be considered for Figure 4.5.

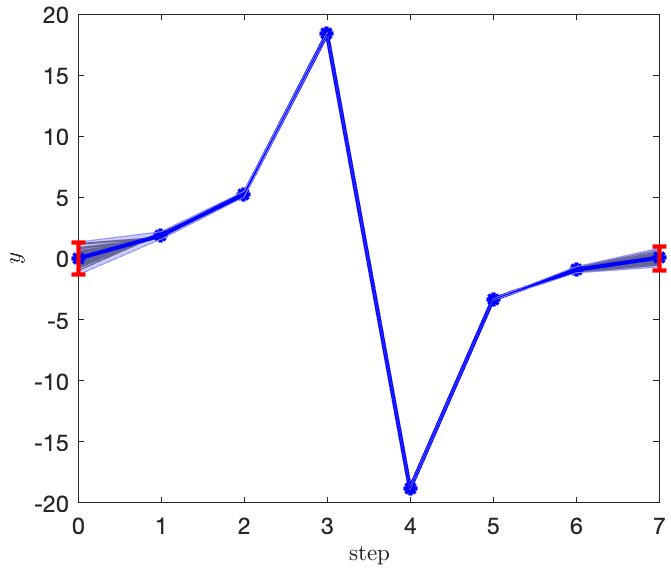


Figure 4.5: Mean and covariance steering for a missile engagement during the end-game

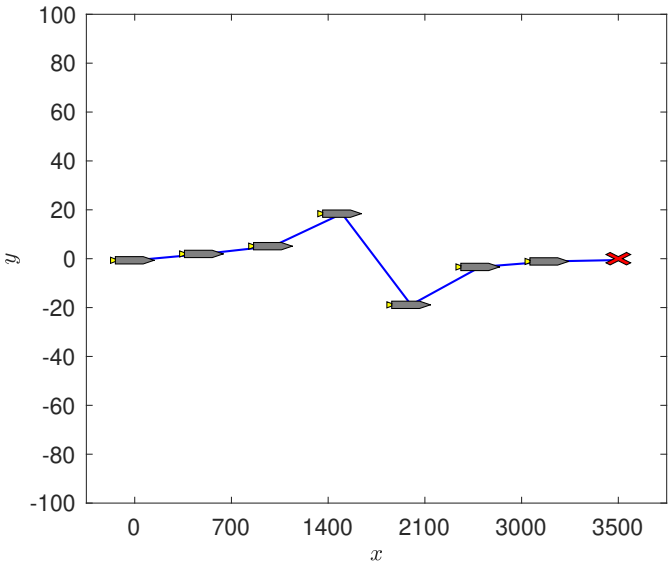


Figure 4.6: Missile's trajectory relative to the target

CHAPTER 5

PURSUIT-EVASION IN STOCHASTIC FLOW FIELDS

In this chapter, we consider two-agent pursuit-evasion problems with both agents traversing a stochastic flow field. It is assumed that both agents have speed constraints, and the pursuer is superior to the evader in terms of its speed capabilities. Initially, a forward reachability analysis is performed while considering only the drift term in the flow field to obtain the nominal trajectories for the agents. We then formulate a discrete-time chance-constrained covariance game about the players' nominal trajectories, which is solved using the standard Gauss-Seidel method to obtain closed-loop control inputs for both players.

5.1 Problem Formulation

Consider a two agent pursuit-evasion scenario in an external stochastic flow field. The dynamics of each agent are given by

$$\mathrm{d}x^i(t) = u^i(t)\mathrm{d}t + G(x^i(t), \mathrm{d}t, \mathrm{d}w^i), \quad x^i(0) = x_0^i, \tag{5.1}$$

where $x^i \in \mathbb{R}^2$, for $i \in \{p, e\}$, denotes the position of an agent (p - pursuer, e - evader) with x_0^i being the i^{th} agent's fixed initial position, known to both players. Here, u^i is the i^{th} agent's control input (velocity) such that $u^i(t) \in \mathbb{R}^2$, and

$$\|u^i(t)\|_2 \leq \mathrm{v}_{\max}^i. \tag{5.2}$$

It is assumed that the pursuer is strictly superior in terms of its speed capabilities compared to the evader, i.e., $\mathrm{v}_{\max}^p > \mathrm{v}_{\max}^e$. The instantaneous dynamic flow field $D(x, \mathrm{d}t, \mathrm{d}w)$ is

assumed to have the form

$$G(x, \mathrm{d}t, \mathrm{d}w) = f(x)\mathrm{d}t + g(x)\mathrm{d}w, \tag{5.3}$$

where $f : \mathbb{R}^2 \to \mathbb{R}^2$ is a position-dependent function, and $g : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$. Here, $w = [w_1\ w_2]^\top$ where w_1 and w_2 are two independent standard Wiener processes. Also, w^p and w^e are assumed to be independent.

In a general pursuit-evasion scenario, the aim of the pursuer is to capture the evader in the shortest time possible, while the evader tries to postpone capture indefinitely. In a deterministic pursuit-evasion game, capture occurs when the Euclidean distance between the agents is less than the capture radius $\varepsilon > 0$. In the proposed formulation, and since the positions of the agents are driven by stochastic processes, capture can only be defined in probabilistic terms. The capture probability at time t is given by

$$\mathcal{C}(t) = \mathbb{P}\{\|x^p(t) - x^e(t)\| \leq \varepsilon\}. \tag{5.4}$$

The goal of this work is to arrive at the control inputs u^i, $i \in \{p, e\}$, that achieve the players' objectives: the pursuer wants to capture the evader in the shortest time possible with high certainty (capture probability); and the evader wants to ensure that the capture probability is as low as possible for all times. To this end, we first obtain the players' nominal trajectories using reachability set analysis while ignoring the disturbance term in the players' dynamics. The system is subsequently linearized along this nominal solution. Then, using the theory of discrete-time linear-quadratic stochastic games, feedback control inputs are constructed that track the trajectory under flow uncertainties while optimizing for the capture probability.

Let $(\hat{x}^i(t), \hat{u}^i(t))$, $i \in \{p, e\}$, for $t \in [0, \hat{T}]$ be the i^{th} player's nominal solution. Here, $\hat{T}$ is the final time of the nominal solution when capture occurs, and $\hat{T} = \infty$ indicate that capture is not possible. The linearized dynamics along the nominal trajectory is given in an

augmented fashion as

$$\mathrm{d}x(t) \approx (u(t) + r(t) + A(t)x(t))\,\mathrm{d}t + D(t)\mathrm{d}w, \tag{5.5}$$

where $x(t) = [x^{p\top}(t), x^{e\top}(t)]^\top$, $u(t) = [u^{p\top}(t), u^{e\top}(t)]^\top$, $w = [w^{p\top}, w^{e\top}]^\top$. Here, $D(t) = \mathrm{blkdiag}(g(x^p(t)), g(x^e(t)))$, $A(t) = \mathrm{blkdiag}(A^p(t), A^e(t))$, where

$$A^i(t) = \frac{\partial f}{\partial x}(\hat{x}^i(t)), \tag{5.6}$$

and $r(t) = [r^{p\top}(t), r^{e\top}(t)]^\top$, where $r^i(t) = f(\hat{x}^i(t)) - A^i(t)\hat{x}^i(t)$.

A discrete-time representation of the dynamics in Equation 5.5 can be expressed as

$$x_{k+1} = A_k x_k + B_k u_k + r_k + D_k w_k, \tag{5.7}$$

where $x_k = x(\tau_k)$, $u_k = u(\tau_k)$, for all $\tau_k = k\hat{T}/N$, $k = \{0, 1, \ldots, N\}$. Assuming a zero-order-hold discretization, we obtain

$$A_k = \Phi(\tau_{k+1}, \tau_k), \tag{5.8a}$$
$$B_k = \int_{\tau_k}^{\tau_{k+1}} \Phi(\tau_{k+1}, \tau)\mathrm{d}\tau, \tag{5.8b}$$
$$r_k = \int_{\tau_k}^{\tau_{k+1}} \Phi(\tau_{k+1}, \tau) r(\tau)\mathrm{d}\tau, \tag{5.8c}$$
$$D_k D_k^\top = \int_{\tau_k}^{\tau_{k+1}} \Phi(\tau_{k+1}, \tau) D(\tau) D^\top(\tau) \Phi^\top(\tau_{k+1}, \tau)\mathrm{d}\tau, \tag{5.8d}$$

where $\Phi(\tau, s)$ is the state transition matrix for the system in Equation 5.7. Using the notation introduced in [94], the system dynamics in Equation 5.7 can be alternatively expressed as

$$X = \mathcal{A}x_0 + \mathcal{B}U + \mathcal{R} + \mathcal{D}W, \tag{5.9}$$

where $X = [x_1^\top,\ x_2^\top, \ldots, x_N^\top]^\top$, $U = [u_0^\top,\ u_1^\top, \ldots, u_{N-1}^\top]^\top$, $W = [w_0^\top,\ w_1^\top, \ldots, w_{N-1}^\top]^\top$, and

$$\mathcal{A} = \begin{bmatrix} A_0 \\ A_1 A_0 \\ \vdots \\ A_{N-1} \ldots A_1 A_0 \end{bmatrix}, \tag{5.10a}$$

$$\mathcal{B} = \begin{bmatrix} B_0 & 0 & \ldots & 0 \\ A_1 B_0 & B_1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{N-1} \ldots A_1 B_0 & A_{N-1} \ldots A_2 B_1 & \ldots & B_{N-1} \end{bmatrix}, \tag{5.10b}$$

$$\mathcal{R} = \begin{bmatrix} r_0 \\ A_1 r_0 + r_1 \\ \vdots \\ A_{N-1} \ldots A_1 r_0 + \cdots + r_{N-1} \end{bmatrix}, \tag{5.10c}$$

$$\mathcal{D} = \begin{bmatrix} D_0 & 0 & \ldots & 0 \\ A_1 D_0 & D_1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{N-1} \ldots A_1 D_0 & A_{N-1} \ldots A_2 D_1 & \ldots & D_{N-1} \end{bmatrix}. \tag{5.10d}$$

The mean and error terms of the augmented position vector are defined as $\bar{X} = \mathbb{E}[X]$ and $\tilde{X} = X - \bar{X}$, respectively.

In order to retrieve the coordinates, and the controls of the agents at each time-step individually from the augmented vectors (X, U), the matrices $E_k = [0_{2\times 2(k-1)}, I_2, 0_{2\times 2(N-k)}]$,

$$E^p = \begin{bmatrix} I_2 & 0_2 & 0_2 & \dots & 0_2 & 0_2 \\ 0_2 & 0_2 & I_2 & \dots & 0_2 & 0_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_2 & 0_2 & 0_2 & \dots & I_2 & 0_2 \end{bmatrix}_{2N \times 4N}, \tag{5.11a}$$

$$E^e = \begin{bmatrix} 0_2 & I_2 & 0_2 & 0_2 & \dots & 0_2 \\ 0_2 & 0_2 & 0_2 & I_2 & \dots & 0_2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_2 & 0_2 & 0_2 & 0_2 & \dots & I_2 \end{bmatrix}_{2N \times 4N}, \tag{5.11b}$$

are introduced. Note that $E^i_k = E_k E^i$, for $i = \{p, e\}$, $1 \leq k \leq N$. Consequently, $E^i X = [x^i_1, \dots, x^i_N]$, and $E^i_k X = x^i_k$.

The relative position vector at time-step k is defined as $x^r_k = x^p_k - x^e_k = E^p_k X - E^e_k X$. The mean and error terms of the relative position vector can be obtained as

$$\begin{aligned} \mathbb{E}[x^r_k] &= \mathbb{E}[E^p_k X - E^e_k X] = (E^p_k - E^e_k)\mathbb{E}[X] \\ &= (E^p_k - E^e_k)\bar{X}, \end{aligned} \tag{5.12}$$

$$\begin{aligned} x^r_k - \mathbb{E}[x^r_k] &= (E^p_k - E^e_k)X - (E^p_k - E^e_k)\bar{X} = (E^p_k - E^e_k)(X - \bar{X}) \\ &= (E^p_k - E^e_k)\tilde{X}. \end{aligned} \tag{5.13}$$

As a result, the covariance of the relative position vector is given by

$$\Sigma^r_k = \mathbb{E}[(E^p_k - E^e_k)\tilde{X}\tilde{X}^\top (E^p_k - E^e_k)^\top]. \tag{5.14}$$

The capture probability at the k^{th} time-step is given by $\mathbb{P}\{\|x^r_k\| \leq \varepsilon\}$. The capture probability can be considered as the payoff function at each time-step, which the pursuer tries to minimize while the evader tries to maximize. An alternative formulation involves

the players optimizing over the minimum value within which the relative distance of the players lie with high probability, given by

$$\bar{\epsilon}_k = \inf\{\epsilon > 0 : \mathbb{P}\{\|x_k^r\| \leq \epsilon\} \geq 1 - \beta\}, \tag{5.15}$$

$1 > \beta > 0$, $1 \leq k \leq N$. The following result provides a lower bound for such an $\bar{\epsilon}_k$ in terms of the mean and the covariance of the relative distance vector x_k^r at the k^{th} time-step.

Theorem 5. *([96]) Let $z \in \mathcal{N}(\mu, \Sigma)$ be an m-dimensional random vector, where $m = 1$ or $m = 2$, let $\sigma = \sqrt{\lambda_{\max}(\Sigma)}$, let $\rho > 0$, and let $1 > \beta > 0$. Then,*

$$\|\mu\| + \sigma\sqrt{2\log\frac{1}{\beta}} \leq \rho \implies \mathbb{P}(\|z\| \leq \rho) \geq 1 - \beta. \tag{5.16}$$

Note that the above result provides a lower bound for any $\epsilon > 0$ that satisfies the condition $\mathbb{P}\{\|x_k^r\| \leq \epsilon\} \geq 1 - \beta$. In this paper, we consider the lower bound of $\bar{\epsilon}_k$, as per Equation 5.16, to be the payoff function that the players try to optimize. To this end, using the result in Theorem 5, we choose to optimize over the mean and covariance of the relative distance vector. From Equation 5.16, it can be observed that by increasing the norm of the mean and/or covariance, the lower bound of $\bar{\epsilon}_k$ can be increased. As a result, the maximizing player can establish guarantees on the minimum relative distance that can be achieved with high probability at every time-step. However, the minimizing player can only hope to minimize $\bar{\epsilon}_k$ by reducing its lower bound, and in this case guarantees on $\bar{\epsilon}_k$ cannot be established.

In the proposed formulation, the mean trajectory is assumed to be essentially driven by the players' controls obtained from the reachability analysis, i.e., the nominal control inputs. Therefore, for the chance-constrained covariance game, the players optimize primarily over the covariance of the relative position vector alone. To this end, we consider

the payoff function, which the pursuer tries to minimize while the evader tries to maximize,

$$J(U) = \|\Sigma^r\|_F^2 = \|\mathbb{E}[(E^p - E^e)\tilde{X}\tilde{X}^\top(E^p - E^e)^\top]\|_F^2, \tag{5.17}$$

subject to the deterministic constraints

$$\|E_k^i(\bar{X} - \hat{X})\|_2 \leq \delta_x^i, \tag{5.18a}$$

$$\|E_k^i(\bar{U} - \hat{U})\|_2 \leq \delta_u^i, \tag{5.18b}$$

for $i = \{p, e\}$. Here, $\hat{X}$ and $\hat{U}$ are concatenated vectors for the nominal solution that are obtained similar to X and U, respectively. In Equation 5.17, it is understood that in the case of the pursuer, by minimizing the norm of the covariance of the augmented relative distance vector, it is minimizing the uncertainty in the relative position at every time-step $1 \leq k \leq N$, and vice versa for the evader. Note that the norm in Equation 5.17 is the Forbenius norm that captures the sum of the squares of the eigenvalues, as opposed to the maximum eigenvalue, suggested by Theorem 5. This is done for the sake of numerical implementation. The constraints in Equation 5.18 ensure that the linearized dynamics in Equation 5.5 remains valid. To account for the control bounds in Equation 5.2, and since a feedback control structure is considered, we also enforce chance constraints at each time-step k of the form

$$\mathbb{P}\{\|E_k^i U\|_2 \leq \mathrm{v}_{\max}^i\} \geq 1 - \beta^i, \quad i = \{p, e\}. \tag{5.19}$$

A reachable set based approach to obtain the nominal trajectories for the players is presented in the next section.

5.2 Reachability Analysis

In this section, the concept of a reachable set is first introduced. To this end, we present some definitions and discuss existing results in the area of reachability set based pursuit-evasion under deterministic flow fields. Finally, a scheme to obtain nominal control inputs of the players using level set methods is presented. In order to obtain the nominal trajectories using reachability analysis, we ignore the disturbance term in Equation 5.3 so that the flow field is deterministic.

The following definitions hold for the case where the agents' dynamics are deterministic.

Definition 9. *An agent's reachable set at time t with the initial state at x_0^i, $\mathcal{R}^i(x_0^i, t)$, $t \geq 0$, is the set of all points that can be reached in time t. The boundary of the reachable set is the reachability front, denoted by $\partial\mathcal{R}^i(x_0^i, t)$.*

Definition 10. *The usable reachable set of the evader $\mathcal{R}_*^e(x_0^e, t)$ is the set of all terminal points of the evader at time t, for which the trajectories do not pass through the reachable set of the pursuer at any time in the interval $[0, t]$. Formally,*

$$\mathcal{R}_*^e(x_0^e, t) = \{x \in \mathbb{R}^2 : x = x^e(t) \text{ and } x^e(\tau) \notin \mathcal{R}^p(x_0^p, \tau),\ \forall\, \tau \in [0, t]\}. \tag{5.20}$$

From Definition Definition 10, it can be observed that the usable reachable set of the evader contains the set of terminal points of the evader's trajectories that are deemed *safe*. If for some time $t_c > 0$, $\mathcal{R}^e(x_0^e, t_c) \subseteq \mathcal{R}^p(x_0^p, t_c)$, then it follows that for every u^e, there exists u^p such that $x^p(t_c) = x^e(t_c)$. In other words, in the deterministic pursuit-evasion scenario, if $\mathcal{R}_*^e(X_0^e, t_c) = \varnothing$, then the capture of the evader is guaranteed at time t_c and vice versa. Consequently, the optimal capture time for pursuit-evasion problems with deterministic flow fields can be established from the following result.

Theorem 6. *([66]) Let $T = \inf\{t \geq 0 : \mathcal{R}_*^e(x_0^e, t) = \varnothing\}$. If $T < \infty$, then capture*

is guaranteed for any time greater than T, whereas the evader can always escape within a time smaller than T. Hence, T is the time to capture if both players play optimally. Furthermore, let x_f denote the location where the evader is captured. Then, we have that $x_f \in \mathcal{X} = \{x \in \mathbb{R}^2 : x = x^e(T)$ *and* $x^e(t) \notin \mathcal{R}^p(x_0^p, \tau),\ \forall\, \tau \in [0, T)\}$

While the above result provides a criterion for the evader's capture based on its usable reachable set, an instantaneous condition that is easier to implement can be stated as follows. For $\mathrm{v}^p_{\max} > \mathrm{v}^e_{\max}$, and assuming the magnitude of the flow field is bounded from above by some suitable constant, we have $\mathcal{R}^e_*(x_0^e, t) = \mathcal{R}^e(x_0^e, t) \setminus \mathcal{R}^p(x_0^p, t)$, for all $t \geq 0$. In such cases, the condition $\mathcal{R}^e_*(x_0^e, t) = \varnothing$ is equivalent to the condition $\mathcal{R}^e(x_0^e, t) \subseteq \mathcal{R}^p(x_0^p, t)$ [66].

The above definitions and results form the crux of the theory related to the reachable set based pursuit-evasion with deterministic equations of motion. The evolution of the reachability front can be traced using, for instance, the level set methods [184]. The reachability front is embedded as a hypersurface in a higher dimension with time as the additional dimension. An implicit representation of the front using the signed distance function is considered in this paper. The signed distance function $\varphi(x)$ with respect to a set $\mathcal{S}$ is defined as

$$\varphi(x) = \begin{cases} \min\limits_{y \in \partial\mathcal{S}} |x - y|, & \text{if } x \notin \mathcal{S}, \\ -\min\limits_{y \in \partial\mathcal{S}} |x - y|, & \text{if } x \in \mathcal{S}. \end{cases} \tag{5.21}$$

For any $c \in \mathbb{R}$, the c-level set of a φ is the set $\{x : \varphi(x) = c\}$. The zero-level set of the signed distance function with respect to an agent's reachable set is expressed as the corresponding reachability front.

Given the signed distance function of an agent's reachable set $\mathcal{R}^i(x_0^i, t)$ as $\phi^i(x, t)$, $i \in \{p, e\}$, the evolution of the corresponding reachability front is governed by the viscosity

solution of the Hamilton-Jacobi equation

$$\frac{\partial \phi^i(x,t)}{\partial t} + \bar{u}^i |\nabla \phi^i(x,t)| + \nabla \phi^i(x,t) f(x) = 0, \tag{5.22}$$

with initial condition $\phi^i(x,0) = |x - x_0^i|$ [185]. Note that $\mathcal{R}^i(x_0^i, t) = \{x \in \mathbb{R}^2 : \phi^i(x,t) \leq 0\}$, and $\partial\mathcal{R}^i(x_0^i, t) = \{x \in \mathbb{R}^2 : \phi^i(x,t) = 0\}$. Once the reachable set of the evader is contained in the pursuer's reachable set at time $\hat{T}$, the nominal trajectories of the players can be obtained in the following manner [185, 66].

From Theorem 6, it can be observed that the terminal point of the evader's nominal trajectory $\hat{x}_f$ is the point in its reachable set that is not covered by the pursuer's reachable set until time $\hat{T}$. Consequently, $\hat{x}_f$ is the point to which the pursuer can drive its nominal trajectory at $\hat{T}$ using its control input $\hat{u}^p$. This point resides on the pursuer's reachability front. When ϕ^i is differentiable, the nominal trajectory of the pursuer can be obtained from the differential equation

$$\frac{\mathrm{d}\hat{x}^p}{\mathrm{d}t} = \hat{\mathrm{v}}^p_{\max} \frac{\nabla \phi^p}{|\nabla \phi^p|} + f(\hat{x}^p), \tag{5.23}$$

and the corresponding optimal control is given by

$$\hat{u}^p = \hat{\mathrm{v}}^p_{\max} \frac{\nabla \phi^p}{|\nabla \phi^p|}. \tag{5.24}$$

Here, $\hat{\mathrm{v}}^p_{\max} \leq \mathrm{v}^p_{\max}$ is the bound on the nominal control input of the pursuer.

Similarly, the evader's nominal trajectory is obtained from the differential equation

$$\frac{\mathrm{d}\hat{x}^e}{\mathrm{d}t} = \hat{\mathrm{v}}^e_{\max} \frac{\nabla \phi^e}{|\nabla \phi^e|} + f(\hat{x}^e), \tag{5.25}$$

and the corresponding optimal control is given by

$$\hat{u}^e = \hat{\mathrm{v}}^e_{\max} \frac{\nabla \phi^e}{|\nabla \phi^e|}, \tag{5.26}$$

where $\hat{\mathrm{v}}^e_{\max} \leq \mathrm{v}^e_{\max}$ is the bound on the nominal control input of the evader. The chance-constrained covariance control problem, introduced in section 5.1, is analyzed and solved using an iterative numerical technique in the following section.

5.3 Covariance Control Game

A linear feedback control structure for the players' control inputs is considered in order to solve the covariance control problem, formulated in section 5.1. Subsequently, the players' inputs are assumed to take the form

$$u_k = v_k + K_k y_k, \tag{5.27}$$

where $K_k \in \mathbb{R}^{4\times 4}$, and $y_k \in \mathbb{R}^4$ is obtained from the difference equation

$$y_{k+1} = A_k y_k + D_k w_k, \quad y_0 = 0. \tag{5.28}$$

In order to calculate y_k, $1 \leq k \leq N-1$, as per Equation 5.28, we assume that each player can observe the instantaneous positions of the players and the control input of their adversary at the previous time-step to evaluate the noise term $D_k w_k$ experienced by the players at an earlier time-step in Equation 5.28. Note that $u_k = [u_k^{p\top}, u_k^{e\top}]^\top$ contains the control inputs of both the pursuer and the evader at time-step k.

Using the matrices introduced in section 5.1, we obtain

$$Y = \mathcal{D} W, \tag{5.29}$$

where $Y = [y_0^\top, \ldots, y_{N-1}^\top]^\top \in \mathbb{R}^{4N}$. Therefore,

$$U = V + KY, \tag{5.30}$$

where $V = [v_0^\top, v_1^\top, \ldots, v_{N-1}^\top]^\top$ and $K = \text{blkdiag}\ (K_0, K_1, \ldots, K_{N-1})$. Substituting Equation 5.30 in Equation 5.9, the mean and the error terms of the augmented state vector X can be obtained as

$$\bar{X} = \mathcal{A}x_0 + \mathcal{B}V + \mathcal{R}, \tag{5.31a}$$

$$\begin{aligned}\tilde{X} &= X - \bar{X} = \mathcal{B}K\mathcal{D}W + \mathcal{D}W \\ &= (I + \mathcal{B}K)\mathcal{D}W,\end{aligned} \tag{5.31b}$$

and for the augmented control vector U, we obtain

$$\bar{U} = \mathbb{E}[U] = V, \quad \tilde{U} = U - \bar{U} = K\mathcal{D}W. \tag{5.32}$$

Subsequently, the covariance matrices are given by

$$\Sigma^y = \mathbb{E}[YY^\top] = \mathcal{D}\mathcal{D}^\top \tag{5.33a}$$

$$\Sigma^x = \mathbb{E}[\tilde{X}\tilde{X}^\top] = (I + \mathcal{B}K)\Sigma^y(I + \mathcal{B}K)^\top \tag{5.33b}$$

$$\begin{aligned}\Sigma^r &= \mathbb{E}[(E^p\tilde{X} - E^e\tilde{X})(E^p\tilde{X} - E^e\tilde{X})^\top] \\ &= (E^p - E^e)\Sigma^x(E^p - E^e)^\top\end{aligned} \tag{5.33c}$$

$$\Sigma^u = \mathbb{E}[\tilde{U}\tilde{U}^\top] = K\Sigma^y K^\top \tag{5.33d}$$

The payoff function in Equation 5.17 can be rewritten in the form

$$\mathcal{J}(V, K) = \|\Sigma^r\|_F^2 = \text{tr}\{(E^p - E^e)^\top(E^p - E^e)\Sigma^x\}. \tag{5.34}$$

The deterministic constraints in Equation 5.18 can be expressed as

$$\|E_k^i(\mathcal{A}x_0 + \mathcal{B}V + \mathcal{R} - \hat{X})\|_2 \leq \delta_x^i, \tag{5.35a}$$

$$\|E_k^i(V - \hat{U})\|_2 \leq \delta_u^i, \tag{5.35b}$$

for $i = \{p, e\}$. Finally, consider the chance constraints on the control inputs of the players given in Equation 5.19. Using Theorem 5, the control chance constraint in Equation 5.19 can be captured using the expression

$$\|E_k^i V\| + \|\Sigma^{y1/2} K^\top E_k^{i\top}\| \sqrt{2 \log \frac{1}{\beta^i}} \leq \mathrm{v}_{\max}^i, \quad i = \{p, e\}. \tag{5.36}$$

The augmented vector V and the matrix K contains the control inputs of both the pursuer and the evader. The pursuer tries to minimize the payoff function in Equation 5.34 by choosing its control inputs $(E^pV, E^pK) \in \mathcal{P}$, while the evader tries to maximize the payoff function by choosing its control inputs $(E^eV, E^eK) \in \mathcal{E}$. the set $\mathcal{P}$ contains all possible tuples $(E^pV, E^pK) \in \mathbb{R}^{2N} \times \mathbb{R}^{2N \times 4N}$ such that the constraints in Equation 5.35 and Equation 5.36 are satisfied for $i = p$ and for all possible $k \in \{0, \ldots, N-1\}$. Similarly, the set $\mathcal{E}$ contains all possible tuples $(E^eV, E^eK) \in \mathbb{R}^{2N} \times \mathbb{R}^{2N \times 4N}$ such that the constraints in Equation 5.35 and Equation 5.36 are satisfied for $i = e$ and for all possible $k \in \{0, \ldots, N-1\}$. Therefore, the proposed stochastic game in section 5.1, given by Equation 5.17-Equation 5.19, is transformed by considering an equivalent payoff function in Equation 5.34, and constraints in Equation 5.35, Equation 5.36. Note that the constraints in Equation 5.35 and Equation 5.36 are orthogonal constraints [153], which are player-specific and are not coupled.

In order to arrive at a saddle point equilibrium for the aforementioned game, assuming one exists, a simple Gauss-Seidel procedure given in Algorithm 3 is considered. For Algorithm 3 to converge to an equilibrium solution for any $V^{(0)}$, $K^{(0)}$, the solution must be stable [180]. The necessary conditions for the existence of a stable equilibrium can be

found in Ref. [180].

Algorithm 3 Gauss–Seidel procedure to obtain saddle points

procedure G-S($V^{(0)}$,$K^{(0)}$)
  for i = 0,1,2,... **do**
    $[E^pV^{(i+1)}, E^pK^{(i+1)}] \leftarrow \underset{(E^pV,E^pK)\in\mathcal{P}}{\arg\min}\ \mathcal{J}(V,K)$ such that $E^eV = E^eV^{(i)}$, $E^eK = E^eK^{(i)}$
    $[E^eV^{(i+1)}, E^eK^{(i+1)}] \leftarrow \underset{(E^eV,E^eK)\in\mathcal{E}}{\arg\max}\ \mathcal{J}(V,K)$ such that $E^pV = E^pV^{(i+1)}$, $E^pK = E^pK^{(i+1)}$
  end for
  return $V^{(i+1)}$, $K^{(i+1)}$
end procedure

5.4 Simulations

In this section, we present simulation results for pursuit-evasion scenarios in realistic flow fields with disturbances. The chance constrained covariance steering game is implemented in MATLAB using its in-built function `fmincon` in conjunction with YALMIP [181]. The convergence criterion for the iterative method is $\|V^{(i+1)} - V^{(i)}\| \leq \epsilon_v$, and $\|K^{(i+1)} - K^{(i)}\| \leq \epsilon_k$. First, a generalized Rankine vortex model is used to generate state-dependent wind field approximations to analyze the performance of the proposed approach on linear flow fields [186]. We then consider a third-order nonlinear flow field that is constructed using orthogonal polynomials [158].

For linear flow fields, the drift part of the wind field is assumed to be of the form

$$f(x) = A(x - x_s), \tag{5.37}$$

and the diffusion part is a constant matrix $g(x) = G$. The simulation parameters are

$$A = \begin{bmatrix} 0.2 & 0.3 \\ -0.15 & 0.1 \end{bmatrix}, \quad x_s = \begin{bmatrix} 10 \\ 10 \end{bmatrix}, \tag{5.38}$$

$x_0^p = [12\ 12]^\top$, $x_0^e = [14\ 14]^\top$, $\hat{T}/N = 0.1$, $\varepsilon = 0.09$, $\mathrm{v}^p_{\max} = 3$, $\mathrm{v}^e_{\max} = 1$, $\hat{\mathrm{v}}^i_{\max} = 0.8\mathrm{v}^i_{\max}$, $\delta^i_u = 0.2\mathrm{v}^i_{\max}$, $\delta^i_x = 0.1$, $\beta^i = 0.01$, $i \in \{p, e\}$, and finally $\epsilon_v = \epsilon_k = 5 \times 10^{-3}$. The diffusion matrix of the flow field is assumed to be of the form $D = \alpha 0.25 I_2$.

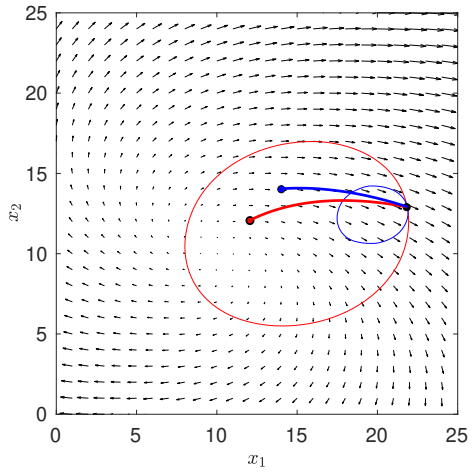


Figure 5.1: Nominal trajectories in the case of a linear flow field

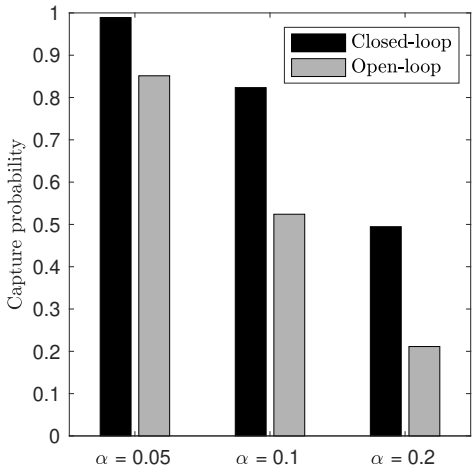
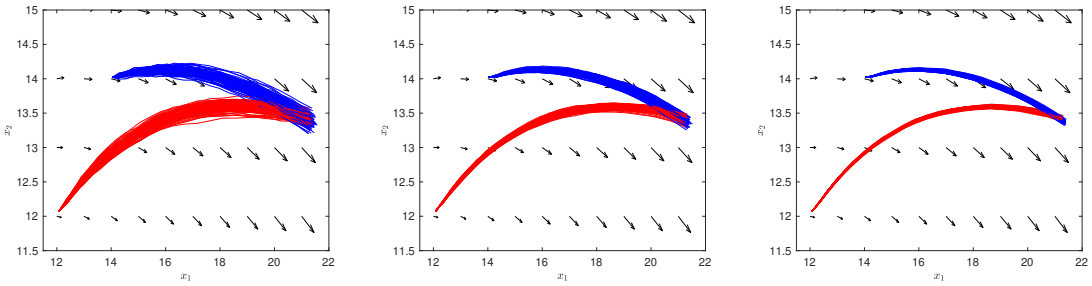


Figure 5.2: Capture probabilities at $\hat{T}$ for different values of α under open-loop and closed-loop control inputs

As explained earlier, the forward reachable sets of the players are first propagated until the pursuer's reachable set fully engulfs the evader's reachable set. The closed curves in the Figure 5.1 are the reachable sets of the players at the final time $\hat{T}$ for the aforementioned



(a) $\alpha = 0.2$ (b) $\alpha = 0.1$ (c) $\alpha = 0.05$

Figure 5.3: Trajectory dispersion of the players in the linear flow field for different α values: Red - pursuer; Blue - evader

simulation parameters. The differential equations in Equation 5.23 and Equation 5.26 are solved backwards in time from the capture point to obtain the nominal trajectories in Figure 5.1 (red - pursuer, blue - evader). Subsequently, the closed-loop trajectories (solid line) are obtained by solving the associated covariance steering game along the nominal (or open-loop) trajectories.

Along the trajectory, the capture probability at each time-step can be obtained by numerically evaluating the integral

$$\mathcal{C}_k = \int_{\|x\| \leq \varepsilon} \frac{1}{2\pi\sqrt{|\Sigma_k^r|}} \exp\left(-\frac{(x-\mu_k^r)^\top \Sigma_k^{-1}(x-\mu_k^r)}{2}\right) \mathrm{d}x, \tag{5.39}$$

where $\mu_k^r = E_k(E^p - E^e)\bar{X}$. Figure 5.2 presents the capture probabilities at the final time ($\hat{T}$) under nominal and optimized control inputs for $\alpha = \{0.05, 0.1, 0.2\}$. Figure 5.3 presents the trajectory dispersion experienced by the players under the closed-loop control for the three α values. It can be observed as α becomes lower, the trajectory dispersion reduces, leading to higher capture probability.

The drift part of the nonlinear flow field, for $x = [x_1,\ x_2]^\top \in \mathbb{R}^2$ is assumed to be of the form

$$f(x) = \begin{bmatrix} a^\top \phi(x_1, x_2) \\ b^\top \phi(x_1, x_2) \end{bmatrix}, \tag{5.40}$$

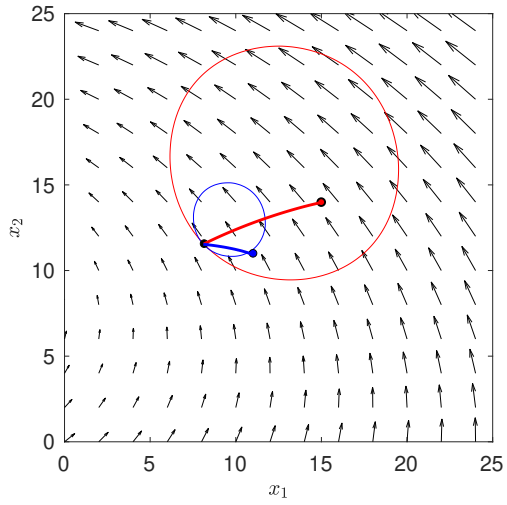
where $\phi(y,z) = [1,\ y,\ z,\ y^2,\ yz,\ z^2,\ y^3,\ y^2z,\ yz^2,\ z^3]^\top$ is the basis of the third-order polynomial vector space. The coefficients are set to

$$a^\top = (1/25) \times [10.8, -0.421, -1.46, -1.78 \times 10^{-3}, 2.42 \times 10^{-3}, 1.07 \times 10^{-4}, \\ -8.61 \times 10^{-7}, 1.17 \times 10^{-7}, -3.03 \times 10^{-5}, -3.32 \times 10^{-8}], \tag{5.41}$$

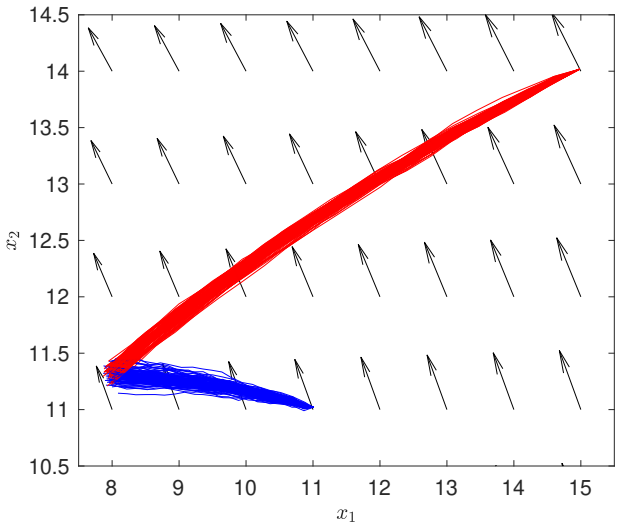
$$b^\top = (1/25) \times [8.67, 0.689, -3.88 \times 10^{-2}, 2.41 \times 10^{-4}, 2.26 \times 10^{-3}, \\ 9.96 \times 10^{-4}, 1.26 \times 10^{-6}, -2.23 \times 10^{-5}, -3.55 \times 10^{-5}, -4.29 \times 10^{-5}]. \tag{5.42}$$

The above nonlinear function represents a single critical point model that was employed to approximate the wind field during Hurricane Hugo [158]. The diffusion part of the nonlinear flow field is consider to be a constant matrix $D = 0.033I_2$. The initial points of the players are chosen to be $x_0^p = [15\ 14]^\top$, $x_0^e = [11\ 11]^\top$, and $\hat{T}/N = 0.15$. The rest of the simulation parameters are same as the ones used for the case of linear flow fields.

The nominal trajectories of the players along with the reachable sets at the final time $\hat{T}$ are shown in Figure 5.4(a). The trajectory dispersion of the players under closed loop control can be seen in Figure 5.4(b). For this simulation, the capture probability at the final time is found to be 46.29% under open-loop control, while it is 82.78% under the closed-loop control. From the simulation results, it can be observed that the proposed closed-loop control strategy for the pursuer is effective in maximizing the capture probability in stochastic flow fields.



(a) Nominal solution



(b) Trajectory dispersion

Figure 5.4: Simulation results in the case of a nonlinear flow field: Red - pursuer; Blue - evader

CHAPTER 6

DESENSITIZED OPTIMAL CONTROL

6.1 Preliminaries

Consider the standard optimal control problem of minimizing the cost

$$\mathcal{J}(u) = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} L(x(t), u(t), t)\,\mathrm{d}t, \tag{6.1}$$

subject to

$$\dot{x}(t) = f(x(t), \rho, u(t), t), \quad x(t_0) = x_0, \tag{6.2a}$$

$$g(x(t), p, t) \leq 0, \tag{6.2b}$$

$$\psi(x(t_f), t_f) = 0, \tag{6.2c}$$

where $t \in [t_0,\, t_f]$ denotes time, with t_0 being the fixed initial time and t_f being the final time, $x(t) \in \mathbb{R}^n$ denotes the state, with x_0 being the fixed state at t_0, and $\rho \in \mathcal{D} \subset \mathbb{R}^\ell$ are model parameters. The control

$$u \in \mathcal{U} = \{u : [t_0, t_f] \to U \text{ is Piecewise Continuous},\ u(t) \in U,\ t \in [t_0, t_f]\} \tag{6.3}$$

with $U \subseteq \mathbb{R}^m$, the set of allowable values of $u(t)$, $\phi : \mathbb{R}^n \times [t_0, t_f] \to \mathbb{R}$ is the terminal cost function, and $L : \mathbb{R}^n \times \mathbb{R}^m \times [t_0, t_f] \to \mathbb{R}$ is the running cost. Here $g : \mathbb{R}^n \times \mathbb{R}^\ell \times [t_0, t_f] \to \mathbb{R}^k$ is a function denoting k state inequality constraints. Finally, $\psi(x(t_f), t_f) = 0$ denotes the terminal condition at time $t = t_f$.

The aforementioned optimal control problem (Equation 6.1 and Equation 6.2) is to be solved by finding the optimal control $u^* \in \mathcal{U}$ that minimizes the cost function in Equa-

tion 6.1, given the constraints in Equation 6.2. The solution defines the optimal state trajectory $x^*(t)$, $t \in [t_0, t_f]$, satisfying $\dot{x}^*(t) = f(x^*(t), \rho, u^*(t), t)$ subject to $x^*(t_0) = x_0$. The system dynamics $f(x, \rho, u, t)$ contains the model parameters, $\rho \in \mathcal{D}$, which are assumed to be constant. In general, the optimal solution $(x^*(t), u^*(t))$ is sensitive to modeling errors and, if changes in the parameter vector ρ occur at any time $t \in [t_0, t_f]$, satisfaction of the constraints in Equation 6.2b or Equation 6.2c is not guaranteed.

In the case of path planning with dynamic obstacles, collision avoidance is of paramount importance. Obtaining safe trajectories under parametric uncertainties in the obstacles' motion is therefore a necessity. In the optimal control framework discussed above, the constraint function in Equation 6.2b can be used to enforce collision avoidance for the path planning problem. Consequently, penalizing a risk measure that captures the possibility of constraint violation under parametric variations may provide the desired safe (e.g., lower chance of constraint violation) trajectories.

With a motivation to minimize the dispersion of the final state of the optimal solution, under parameter uncertainties, Seywald and Kumar constructed an augmented cost function using sensitivity matrices [111]. It should be noted that the sensitivity matrix, from Refs. [111, 113], differs from the standard sensitivity functions that are defined in Ref. [164].

The approach goes as follows. First, the parameters of interest and the corresponding entries in sensitivity matrix are elevated to states, and the augmented state $[\tilde{x}^\top \ (\mathrm{vec}\ \Lambda)^\top]^\top$, where $\tilde{x} = [x \ \ \rho]^\top$, along with the corresponding dynamics and initial conditions are derived as follows

$$\dot{\tilde{x}} = \tilde{f}(\tilde{x}, u, t) = [f^\top(x, \rho, u, t) \ \ 0_{1\times\ell}]^\top, \quad \tilde{x}(t_0) = [x_0^\top \ \ \hat{\rho}^\top]^\top, \tag{6.4}$$

and

$$\dot{\Lambda}(t|t_0, \tilde{x}_0) = \frac{\partial \tilde{f}}{\partial \tilde{x}}(x, \rho, u, t), \quad \Lambda(t_0|t_0, \tilde{x}_0) = I_{(n+\ell)}, \tag{6.5}$$

where $\rho(t) \in \mathcal{D}$ denotes the ℓ parameters of interest and $\hat{\rho}$ is the nominal value of these parameters, and $\Lambda(t|t_0, x_0)$ represents the sensitivity of the vector $\tilde{x}(t)$ at time t with respect to perturbations in the initial state vector $\tilde{x}(t_0)$. That is,

$$\Lambda(t|t_0, x_0) = \frac{\partial \tilde{x}(t)}{\partial \tilde{x}(t_0)}. \tag{6.6}$$

The augmented cost function, given in Equation 6.7 below, is then minimized to obtain an optimal solution with final state being "desensitized" with respect to the parameter variations.

$$\mathcal{J}_a(u) = \mathcal{J}(u) + \int_{t_0}^{t_f} \|\text{vec}\ \left(\Lambda(t_f|t_0, \tilde{x}_0)\Lambda(t|t_0, \tilde{x}_0)^{-1}\right)\|^2_{Q(t)}\, \mathrm{d}t, \tag{6.7}$$

with $Q(t) \geq 0,$ for all $t \geq 0$. Note that the sensitivity matrix of Seywald in Equation 6.6 is a state transition matrix and its properties are exploited to construct the sensitivity of the final state with respect to the variations in the state at time $t \in [0, t_f]$, which is then plugged into the running cost. This is elaborated upon in [111]. However, this approach requires propagating the original states, the targeted parameters, and the elements in the sensitivity matrix, resulting to a total of $(n + \ell)^2 + n + \ell$ number of states.

In this chapter, we first employ the traditional sensitivity functions to develop an alternative scheme for optimal trajectory/state desensitization with respect to parameter variations with improved computational efficiency. Two versions of the Zermelo's navigation problem are employed to demonstrate the efficacy of the proposed schemes for trajectory and final state desensitization. We then develop a scheme to generate constraint desensitized trajectories by penalizing a risk measure that is defined using sensitivity functions. Finally, we prove that the co-states indeed capture sensitivity of the *cost-to-go* function

with respect to perturbations in the state given *any* prescribed control law $u(t)$, not just the optimal one. Using this fact, a new approach to solve the desensitized optimal control problem is presented.

6.2 Sensitivity Functions

Consider the dynamics in Equation 6.2a, and assume variations in the model parameters $\rho \in \mathcal{P}$, with $\rho = \hat{\rho}$ being the nominal value of the parameter vector. Furthermore, assume that $f(x, \rho, u, t)$ is continuous in (x, ρ, u, t), and continuously differentiable with respect to x and ρ for all $(x, \rho, u, t) \in \mathbb{R}^n \times \mathcal{D} \times U \times [t_0, t_f]$. The solution to the differential equation from the initial condition x_0 with control input $u \in \mathcal{U}$ is given by

$$x(\rho, t) = x_0 + \int_{t_0}^{t} f(x(\rho, \tau), \rho, u(\tau), \tau)\, \mathrm{d}\tau. \tag{6.8}$$

Since $f(x, \rho, u, t)$ is differentiable with respect to ρ, it follows that

$$\frac{\partial x(\rho, t)}{\partial \rho} = \int_{t_0}^{t} \left[\frac{\partial f(x(\rho, \tau), \rho, u(\tau), \tau)}{\partial x} \frac{\partial x(\rho, \tau)}{\partial \rho} + \frac{\partial f(x(\rho, \tau), \rho, u(\tau), \tau)}{\partial \rho} \right] \mathrm{d}\tau. \tag{6.9}$$

Taking the derivative with respect to t, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{\partial x(\rho, t)}{\partial \rho} \right] = \frac{\partial f(x, \rho, u(t), t)}{\partial x} \frac{\partial x(\rho, t)}{\partial \rho} + \frac{\partial f(x, \rho, u(t), t)}{\partial \rho}. \tag{6.10}$$

Evaluating Equation 6.10 at the nominal conditions ($\rho = \hat{\rho}$), the dynamics for the *parameter sensitivity function* $S : [0, \infty) \to \mathbb{R}^{n \times \ell}$

$$S(t) = \left. \frac{\partial x(\rho, t)}{\partial \rho} \right|_{x = x(\hat{\rho}, t)} \tag{6.11}$$

can be obtained as

$$\dot{S}(t) = A(t)S(t) + B(t), \quad S(t_0) = 0_{n\times\ell}, \tag{6.12}$$

where

$$A(t) = \left.\frac{\partial f(x,\rho,u(t),t)}{\partial x}\right|_{x=x(\hat{\rho},t),\,\rho=\hat{\rho}}, \tag{6.13a}$$

$$B(t) = \left.\frac{\partial f(x,\rho,u(t),t)}{\partial \rho}\right|_{x=x(\hat{\rho},t),\,\rho=\hat{\rho}}. \tag{6.13b}$$

Since the initial state is given (fixed), the initial condition for the sensitivity function is the zero matrix, and Equation 6.12 is called the *sensitivity equation* in the literature [164]. To compute the sensitivity function over time, the state x has to be propagated using the dynamics in Equation 6.2a under nominal conditions,

$$\dot{x} = f(x,\hat{\rho},u,t), \quad x(t_0) = x_0. \tag{6.14}$$

From the properties of continuous dependence with respect to the parameters and the differentiability of solutions of ordinary differential equations, and for sufficiently small variations in $\hat{\rho}$, the solution $x(\rho,t)$ can be approximated by

$$x(p,t) \approx x(\hat{\rho},t) + S(t)(\rho - \hat{\rho}). \tag{6.15}$$

This is a first-order approximation of $x(\rho,t)$ about the nominal solution $x(\hat{\rho},t)$.

Remark 3. *The difference between the sensitivity function and the sensitivity matrix lies in the fact that the former cannot accommodate time-varying parameters which is evident in (Equation 6.15). On the other hand, sensitivity matrices can be used to investigate variation in the state $x(t)$ (at time t) with respect to a variation in the parameter $\rho(t')$ (at any other time $t_0 \leq t' \leq t$) i.e., the variations can be different along the trajectory and the*

parameter can be time-varying. In this regard, though sensitivity matrices can handle more complexity, this comes with an added cost in terms of computation. In addition, for most problems, the model parameters are some constants whose values are prone to changes from the nominal.

6.3 Optimal Trajectory Desensitization

Trajectory desensitization allows one to find a robust path that is relatively immune to parameter variations, thereby mitigating the requirements on the feedback controller that may be used to track the obtained optimal trajectory. Revisiting the standard optimal control problem discussed in section 6.1 (Equation 6.1 and Equation 6.2), consider minimizing the augmented cost

$$\mathcal{J}_a(u) = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} \Big(L(x(t), u(t), t) + \|\text{vec } S(t)\|^2_{Q(t)} \Big) \mathrm{d}t, \tag{6.16}$$

with an augmented state $[x^\top \ (\text{vec } S)^\top]^\top$, whose dynamics is obtained from Equation 6.14 and (Equation 6.12), while imposing the terminal condition (Equation 6.2c). The augmented cost function in Equation 6.16 minimizes the original cost function in Equation 6.1, while penalizing the sensitivity of the state with respect to the parameters along the optimal trajectory. The weighting factor for the sensitivity cost, $Q(t)$, can be tuned to balance between minimizing the original cost and the sensitivity cost.

If one is interested only in the variation of the final state (or state at a particular time instant $t > t_0$) with respect to parameter variations, then the corresponding sensitivity terms alone can be penalized by adding an extra term to the terminal cost of the original cost function as follows.

$$\mathcal{J}_a(u) = \phi(x(t_f), t_f) + \|\text{vec } S(t_f)\|^2_Q + \int_{t_0}^{t_f} L(x(t), u(t), t) \, \mathrm{d}t, \tag{6.17}$$

where $Q \geq 0$ is the weighing factor for the terminal sensitivity cost. Final state desensiti-

zation is critical especially among problems involving unmanned vehicles where the goal state has to be reached precisely under external disturbances. An alternative to this approach is to directly constrain the sensitivity function at the final time by adding additional state constraints of the form

$$\tilde{\psi}(S(t_f), t_f) = 0. \tag{6.18}$$

6.3.1 Numerical Simulations

Consider the Zermelo's problem [111] with currents parallel to the shore (assumed to be the x_1-axis) as a function of x_2 such that

$$\mathrm{v}_{\text{current}} = \rho x_2, \tag{6.19}$$

where ρ is a parameter which is uncertain and its nominal value is $\hat{\rho}$. The dynamics for a boat traveling in the currents can be written as

$$\dot{x}_1 = \cos(u) + \rho x_2, \tag{6.20}$$

$$\dot{x}_2 = \sin(u), \tag{6.21}$$

where u is its heading control, $u \in \mathcal{U} = \{\text{PWC}, u(t) \in (-\pi, \pi],\ \forall\, t \in [0, t_f]\}$, for some $t_f > 0$. In this example, the cost function that has to be minimized is expressed in Mayer form as

$$\mathcal{J}(u) = -x_1(t_f), \tag{6.22}$$

and the boundary conditions are

$$x_1(0) = 0, \quad x_2(0) = 0, \quad x_2(t_f) = 0. \tag{6.23}$$

For this example, we let $t_f = 1$, and $\hat{\rho} = 10$. Apart from maximizing the length that the boat can traverse along the shore while meeting the boundary conditions, the optimal trajectory has to be desensitized with respect to the variations in ρ. The goal can be facilitated by obtaining the sensitivity equation under nominal conditions ($\hat{\rho} = 10$), using Equation 6.12,

$$\dot{S} = \begin{bmatrix} \dot{S}_1 \\ \dot{S}_2 \end{bmatrix} = \begin{bmatrix} \hat{\rho} S_2 + x_2 \\ 0 \end{bmatrix}, \quad S(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{6.24}$$

where

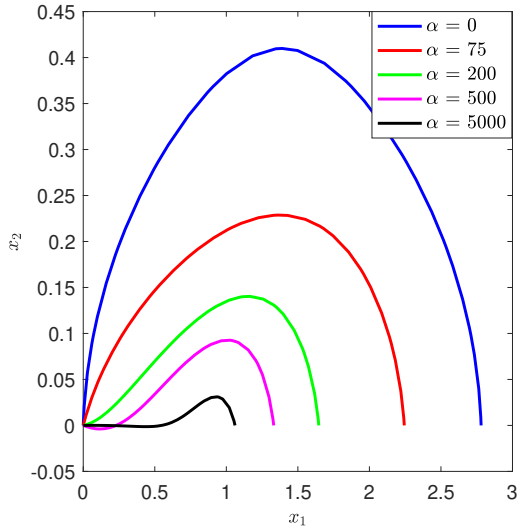
$$S_i(t) = \left. \frac{\partial x_i(\rho, t)}{\partial \rho} \right|_{x = x(\hat{\rho}, t)}, \; i = 1, 2, \tag{6.25}$$

and then by penalizing the terms S_1 and S_2 in an augmented cost function, as shown in Equation 6.16. Note that $S_2(t) = 0, \;$ for all $t \geq 0$, which means that the state x_2 is not affected by the uncertainty of the currents along a nominal trajectory. Therefore, in this particular example, just penalizing S_1 is sufficient. The augmented cost function can then be written as

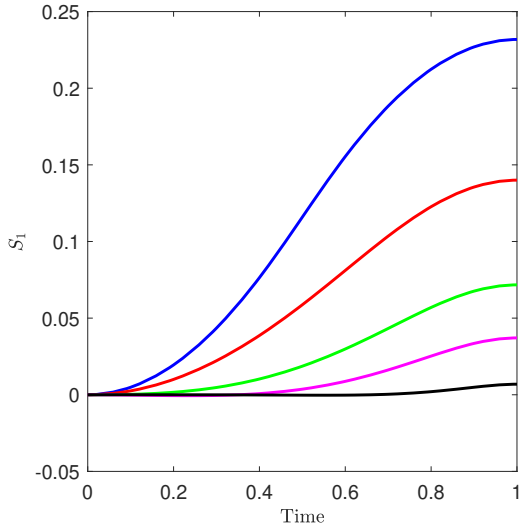
$$\mathcal{J}_a(u) = -x_1(t_f) + \int_{t_0}^{t_f} \alpha S_1^2(t) \, \mathrm{d}t. \tag{6.26}$$

The weighting factor α (which is a constant in this case) is chosen by the designer. For $\alpha = 0$, the optimal solution for the original cost function (Equation 6.22) is obtained. Several test cases were run for $\alpha = \{0, 75, 200, 500, 5000\}$, using GPOPS-II [170], and the results are shown in Figure 6.1 and Figure 6.2.

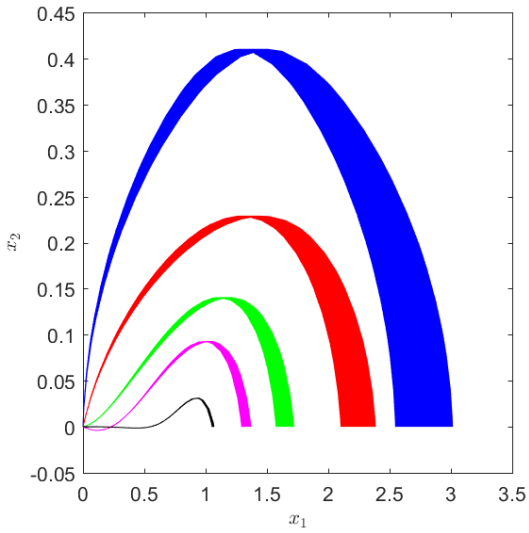
The levels of desensitization can be clearly observed in Figure 6.1. As the value of α increases, the trajectory becomes more conservative by staying closer to the shore, while trying to maximize $x_1(t_f)$; see Figure 6.1(a). At the same time, the magnitude of the sensitivity of x_1 with respect to ρ (that is, S_1) also decreases along the trajectory which is observed in Figure 6.1(b). Monte-Carlo simulations, shown in Figure 6.1(c), illustrate the idea of trajectory desensitization. It can be seen that with high weight on α, the trajectories



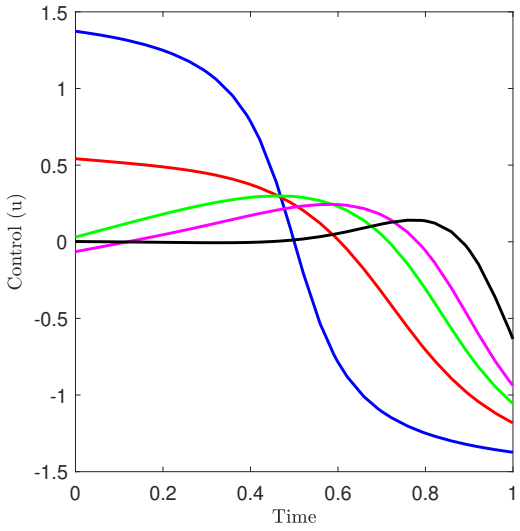
(a) Optimal trajectories



(b) Sensitivity of x_1 with respect to ρ (S_1)



(c) Monte-Carlo simulations with open-loop control



(d) Optimal Control (u^*)

Figure 6.1: Results obtained for the Zermelo's path optimization problem with trajectory desensitization

obtained using open-loop control under parameter variations stay closer to the corresponding optimal trajectory. For the Monte-Carlo simulations that are shown in Figure 6.1(c) and in the rest of the paper, the corresponding parameter ρ was chosen randomly between $\pm 10\%$ about its nominal value and it is kept constant for one Monte-Carlo run.

For the simulations with feedback controller in Figure 6.2, a linear quadratic regulator is constructed by minimizing the cost

$$\mathcal{J}_f = \frac{1}{2}\|\Delta x(t_f)\|^2 + \int_0^{t_f} \Delta u^2(t)\,\mathrm{d}t, \tag{6.27}$$

where $\Delta x(t) = x(t) - x^*(t)$, $\Delta u(t) = u(t) - u^*(t)$, and by linearizing the dynamics along the reference trajectory as

$$\Delta \dot{x} = \mathcal{A}(t)\Delta x + \mathcal{B}(t)\Delta u, \quad \Delta x(t_0) = 0_{n\times 1}, \tag{6.28}$$

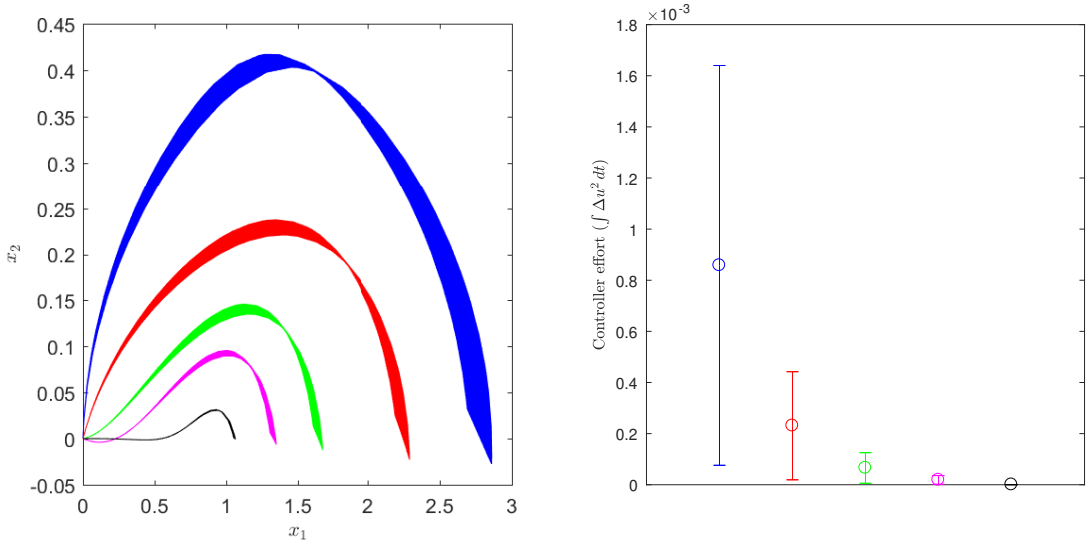
where

$$\mathcal{A}(t) = \left.\frac{\partial f(x, \hat{\rho}, u, t)}{\partial x}\right|_{x=x^*(t),\, u=u^*(t)}, \tag{6.29a}$$

$$\mathcal{B}(t) = \left.\frac{\partial f(x, \hat{\rho}, u, t)}{\partial u}\right|_{x=x^*(t),\, u=u^*(t)}. \tag{6.29b}$$

Consequently, the control effort required to track the optimal trajectory (for some α) under parameter variations, using the feedback controller, reduces as the value of α increases, as can be observed in Figure 6.2(b). These results further corroborate the proposed approach for desensitized optimal control that deals with the trade-off between optimality and tracking effort using a feedback controller.

Next, we consider the time-optimal Zermelo's problem. The dynamics presented in the previous subsection (Equation 6.21), and the initial conditions (Equation 6.23) are retained. The vehicle, starting from the origin, has to reach a point along the shore $(2, 0)$ in minimum time. However, the parameter ρ is uncertain, and the goal state $(2, 0)$ has to be



(a) Monte-Carlo Simulations with feedback control

(b) Trends of feedback controller effort obtained from Monte-Carlo simulations

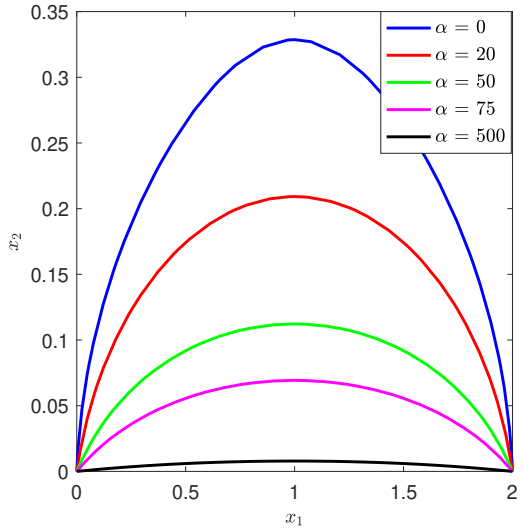
Figure 6.2: Results for trajectory desensitization with feedback control

reached as accurately as possible i.e., the sensitivity of final state with respect to the parameter variation is of concern. For this purpose, the augmented cost function, equivalent to Equation 6.17, can be constructed as

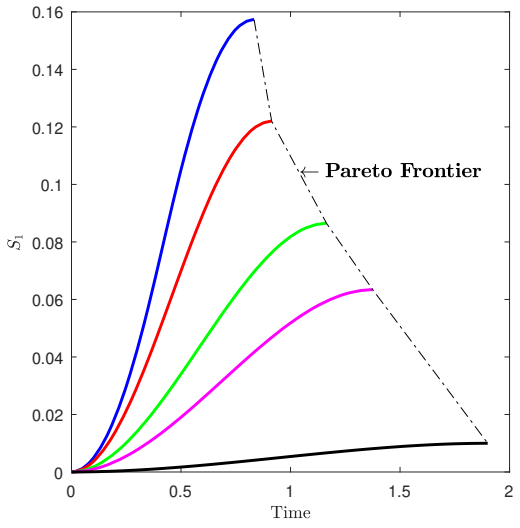
$$\mathcal{J}_a(u) = t_f + \alpha S_1^2(t_f). \tag{6.30}$$

The sensitivity equation (Equation 6.24) too remains the same for this example, as it only depends on the dynamics of the problem and the targeted parameters, but not on the original cost function. The results obtained from the simulations for $\alpha = \{0, 25, 75, 1000\}$ while minimizing the cost function in Equation 6.30 are presented in Figure 6.3.

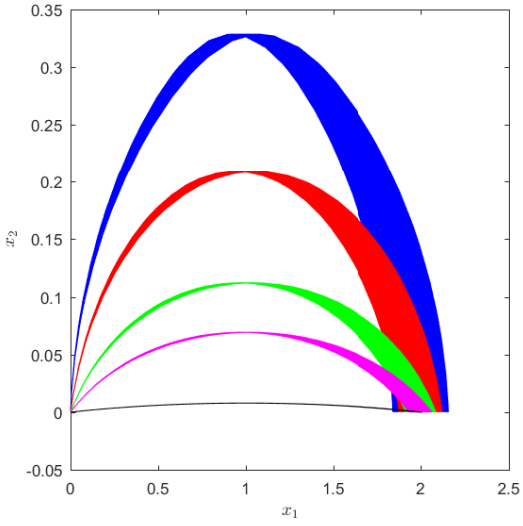
All optimal trajectories meet the boundary conditions at the initial and final times, but they differ in their final state sensitivity with respect to the uncertain parameter. A clear trade-off between the time taken to reach the goal state (t_f) and final state sensitivity ($S_1(t_f)$) can be observed from Figure 6.3(b) and Figure 6.3(c). A Pareto frontier can be drawn to quantitatively establish the trade-off between $S_1(t_f)$ and t_f, which is also shown



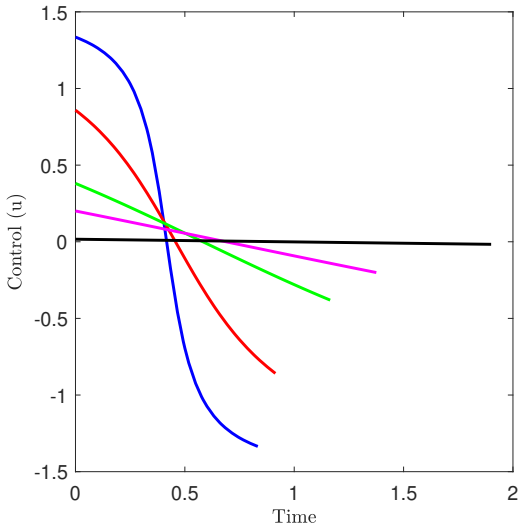
(a) Optimal trajectories



(b) Sensitivity of x_1 with respect to ρ (S_1)



(c) Monte-Carlo simulations with open-loop control



(d) Optimal Control (u^*)

Figure 6.3: Results obtained for the Zermelo's path optimization problem with final state desensitization

in Figure 6.3(b). The Monte-Carlo simulations, in which the parameter is again randomly varied between $\pm 10\%$ about the nominal value, further strengthen the claim that the dispersion in the final state can be reduced at the cost of the time it takes to reach the goal state. Note the similarity in the trends of the sensitivity term S_1 with the trajectory desensitization by comparing Figure 6.1(b) and Figure 6.3(b). The reason behind the similarity could be because of the fact that since $S_2(t) = 0,$ for all $t \geq 0$, $\dot{S}_1 = x_2$. This implies that S_1 is almost always increasing along the trajectory, as the slope is mostly positive, and penalizing S_1 at the final state indirectly penalizes its slope at every point along the trajectory, thus desensitizing the entire trajectory in this example. The results can also be compared with Seywald's in Ref. [111]. The effect of desensitization remains the same, while achieving the goal with fewer number of states in our approach. The number of states in the augmented model is 4 with the proposed approach, whereas it is 21 in the case of Seywald's approach.

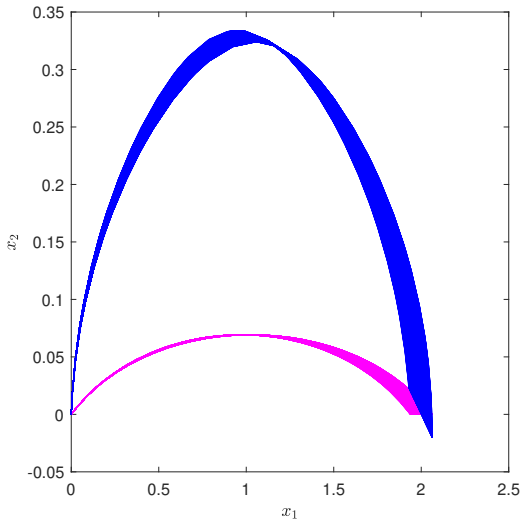
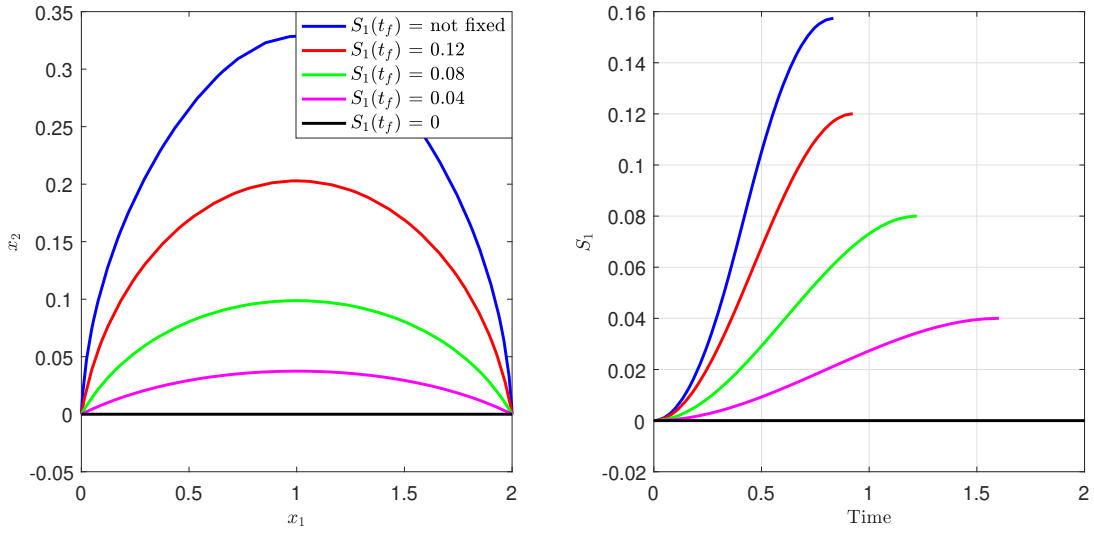


Figure 6.4: Monte-Carlo simulations for non-desensitized trajectory with feedback controller and desensitized trajectory with open-loop

In Figure 6.4, a comparison between the non-desensitized trajectory with feedback controller and a desensitized trajectory ($\alpha = 75$, the color code remains the same as in Figure 6.3) with open-loop control is made again Monte-Carlo simulations. The feedback



(a) Optimal trajectories

(b) Sensitivity of x_1 with respect to ρ (S_1)

Figure 6.5: Results obtained for the Zermelo's path optimization problem - Sensitivity of the final state with respect to the uncertain parameter is fixed

controller design remains the same as expressed in Equation 6.27-Equation 6.29. It can be observed that the dispersion in x_1 at final time ($x_1(t_f)$) is almost the same in both cases. However, the energy of the overall control signal, $\int_{t_0}^{t_f}(u(t) + \Delta u(t))^2 \mathrm{d}t = 0.08$, on an average is significantly higher in the non-desensitized case, as opposed to 0.02 in the case of desensitization (open-loop). This indicates that by employing desensitization to work, the dispersion in the final state can be reduced with lesser control effort by compromising on the original cost, which is travel time in this example.

The results obtained for Zermelo's problem while directly fixing the value of S_1 at final time for different cases, $S_1(t_f) = \{\text{not fixed}, 0.12, 0.08, 0.04, 0\}$, are presented in Figure 6.5. It is interesting to observe that the optimal trajectory for the case $S_1(t_f) = 0$ goes along the shore without venturing into the currents, as to be expected from the physics of this problem.

The proposed scheme outperforms the approach presented in Ref. [111] in terms of the computation complexity with just $n + n\ell$ number of states, as opposed to $(n + \ell)^2 + n + \ell$ number of states in the latter approach. Also, the freedom in formulating desensitized opti-

mal control schemes, especially for trajectory desensitization, is significant with sensitivity functions. However, given the dynamics and the constraints (Equation 6.1-Equation 6.2), there exists a set of trajectories, and each trajectory has an associated sensitivity cost $\left(\int_{t_0}^{t_f} \|\text{vec } S(t)\|^2_{Q(t)} \mathrm{d}t\right)$. Given the original cost function (Equation 6.1), the optimal trajectory that minimizes the cost function also has an associated sensitivity cost. In the proposed approached with an augmented cost function, by penalizing the sensitivity cost, one expects to find another trajectory that is less sensitive, while satisfying all the constraints. Such a trajectory, however, may not exist. There is an implicit assumption here, namely, that there exists a path that is less sensitive to parametric variations compared to the optimal one, given the original cost function along with state and control constraints. In some problems, it could be the case where the solution to the original optimal control problem can no longer be desensitized. Analysis on "desensitizability" of an optimal control problem is a potential direction for future work.

Similarly, in the case of final state desensitization where the sensitivities at the final time are fixed, it is assumed that the control u is able to drive the augmented state, with the sensitivity terms added to the original state vector x, given the constraint in the original optimal control problem. In that case, in order to be able to desensitize a given system, there is an implicit assumption that the system with the augmented state $[x^\top \ (\text{vec } S)^\top]^\top$ is controllable. Otherwise, the control input u may not have enough authority to drive the additional states introduced via the sensitivity function. In such cases, an additional feedback term is the only available option to handle the parameter variations.

6.4 Constraint Desensitized Path Planning

For the optimal control problem (Equation 6.1)-(Equation 6.2), assuming the constraint function $g(x, \rho, t)$ is a smooth function in x, a naïve approach to obtain conservative trajectories to address constraint violation under parametric uncertainties would be to penalize

the *constraint sensitivity* matrix, defined as

$$\begin{aligned} S_g(t) &= \left.\frac{\partial g(x(\rho,t),\rho,t)}{\partial \rho}\right|_{\rho=\hat{\rho}} \\ &= \left(\frac{\partial g(x,\rho,t)}{\partial x}S(t) + \frac{\partial g(x,\rho,t)}{\partial \rho}\right)\Bigg|_{x=x(\hat{\rho},t),\rho=\hat{\rho}}, \end{aligned} \tag{6.31}$$

by constructing the augmented cost

$$\mathcal{J}_g(u) = \mathcal{J}(u) + \int_{t_0}^{t_f} \|\text{vec } S_g(t)\|_Q^2 \,\mathrm{d}t, \tag{6.32}$$

where $Q \geq 0$.

By minimizing the cost in Equation 6.32, one attempts to minimize the variation in the constraint value with respect to variations in the parameter for all times. However, it is clear that the variation in the constraint value when the system is far from the constraint boundary is not as important as when the system is close to the constraint boundary. For example, in the path planning problem with a dynamic obstacle, a larger variation in the constraint value may be acceptable when the agent is far from the obstacle, but a collision may result due to even relatively small variations when the agent is near the constraint boundary. Weighting the sensitivity of the constraint value equally in both cases using a running cost function may therefore lead to solutions that are highly sensitive near the constraint boundary.

To account for the fact that the constraint variations are more likely to cause constraint violations when the system is closer to the constraint boundary, we introduce a relevance function $r : \mathbb{R} \to [0, \infty)$ of the form

$$r(w) = \begin{cases} \tilde{r}(w), & \text{if } w \leq 0, \\ \tilde{r}(0), & \text{if } w > 0, \end{cases} \tag{6.33}$$

where $\tilde{r} : \mathbb{R} \to [0, \infty)$ is a continuous function that is monotonically increasing over the

interval $(-\infty, 0]$, that is, $\tilde{r}(w_1) \geq \tilde{r}(w_2)$, if $w_1 > w_2$ for all $w_1, w_2 \leq 0$. Examples of $\tilde{r}(w)$ include e^{-w^2} (Gaussian), $\max(0, 1 - |w|)$ (Hat function), $1/(1 + w^2)$, etc.

Next, we define the *relevant constraint sensitivity* (RCS) matrix $S_r : [0, \infty) \to \mathbb{R}^{k \times \ell}$ as

$$S_r(t) = RS_g(t), \tag{6.34}$$

where $R = \text{diag}\left(r(g_1(x(p,t),p,t)), \ldots, r(g_k(x(p,t),p,t))\right)$. Henceforth for the purpose of the analysis, and unless stated otherwise, the derivative of the logistic function $s(w) = 1/(1 + e^{-w})$ is chosen as the candidate relevance function, that is,

$$\tilde{r}(w) = s(w)(1 - s(w)). \tag{6.35}$$

Note that the derivative of the logistic function has a symmetric "bell-shape" with the maximum at $w = 0$, and decaying tails. The relevance function allows one to penalize sensitivities according to their relevance with respect to potential constraint violation. The impact of the choice of the relevance function on constraint desensitization is discussed later. The sensitivity matrix S_r captures the idea of giving more importance to variations near the constraint boundary.

Finally, we propose to solve the optimal control problem with the augmented cost function

$$\mathcal{J}_a(u) = \mathcal{J}(u) + \int_{t_0}^{t_f} \|\text{vec } S_r(t)\|_Q^2 \, \mathrm{d}t, \tag{6.36}$$

the dynamics in Equation 6.14 and Equation 6.12, and the constraints in Equation 6.2 to construct trajectories that address constraint violation under parametric uncertainties. Hereafter, the term $\int_{t_0}^{t_f} \|\text{vec } S_r(t)\|_Q^2 \, \mathrm{d}t$ in Equation 6.36 will be referred to as the RCS cost.

We now apply the proposed approach on simple test examples to analyze the optimal trajectories obtained by penalizing RCS. First, we analyze the claim of penalizing RCS over

constraint sensitivity using a 2D path planning problem involving a dynamic obstacle with uncertainty in its speed. Subsequently, the effect of various constraint forms that represent the collision avoidance condition, chosen from a set of valid ones, is studied. We then stress upon the need to select an appropriate constraint function to construct RCS using the car vs. train problem, and finally, the trade-off studies with multiple obstacles are presented. The videos demonstrating the optimal trajectories for the example problems discussed in this section can be found on the web[1].

6.4.1 2D Time-Optimal Problem

Consider the following 2D time-optimal path planning problem with the agent dynamics and initial conditions

$$\dot{x}(t) = \mathrm{v}\cos(u(t)), \quad x(0) = a_0, \tag{6.37a}$$

$$\dot{y}(t) = \mathrm{v}\sin(u(t)), \quad y(0) = b_0, \tag{6.37b}$$

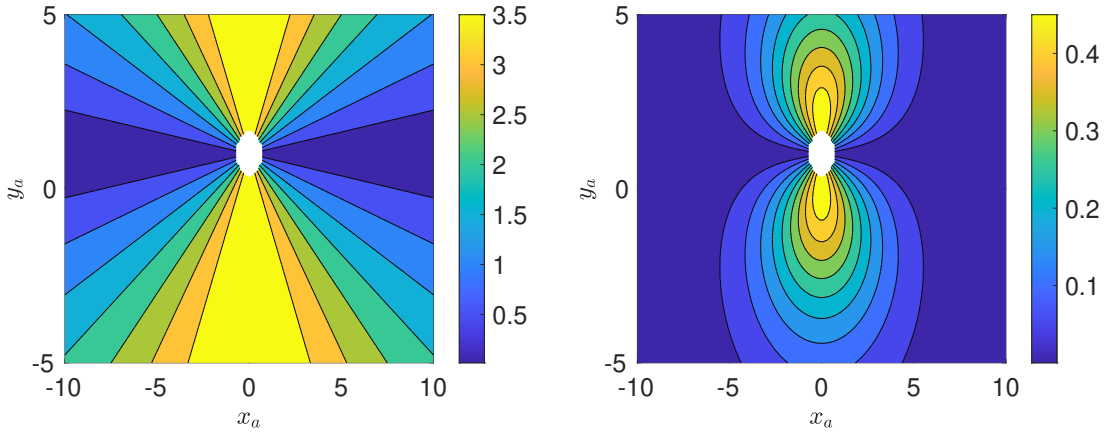
where (x, y) denotes the agent's position, v is the agent's speed, and $u(t) \in [0, 2\pi)$ is the agent's heading (control). The agent intends to reach (a_f, b_f) in minimum time, while avoiding a dynamic circular obstacle that is moving parallel to the y-axis with a constant speed ρ, and dynamics given by

$$\dot{z}(t) = \rho, \quad z(0) = d, \tag{6.38}$$

where (c, z) denotes the obstacle's position. With $[x, y, z]^\top$ as the state vector, the constraint for collision avoidance can be expressed as

$$g = r_o - \left[(x-c)^2 + (y-z)^2\right]^{1/2} \leq 0, \tag{6.39}$$

[1] https://youtu.be/zCvuIQSMzlw



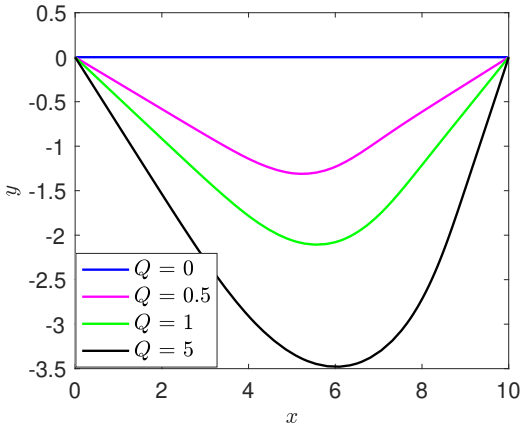
(a) $|S_g|$ (b) $|S_r|$

Figure 6.6: Absolute values of S_g and S_r as a function of the agent's position

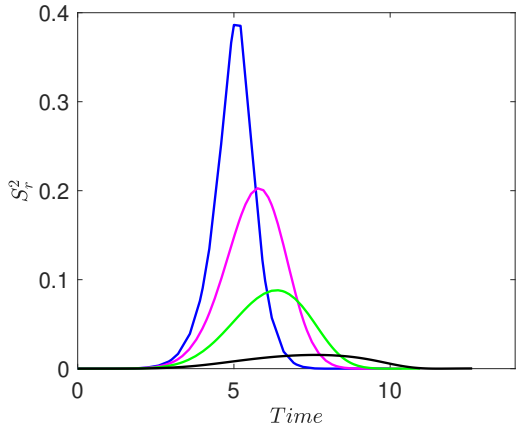
where r_o is the safe distance. For this problem, we assume that the obstacle's speed ρ is the uncertain parameter. Henceforth, the time and parameter dependency of the elements in the state vector are dropped for brevity. Since the problem has simple dynamics, a closed form expression to the sensitivity of the constraint function $g(x)$ with respect to the uncertain parameter ρ is given by

$$S_g(t) = \frac{\partial g}{\partial \rho} = -\frac{(y-z)t}{[(x-c)^2 + (y-z)^2]^{1/2}}. \tag{6.40}$$

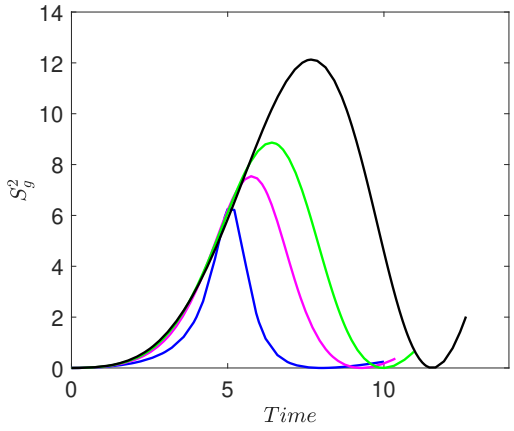
The effect of incorporating the relevance function into the proposed scheme is analyzed by comparing the constraint sensitivity and RCS. In this regard, the obstacle's position is fixed at $(0, 1)$ with $r_o = 0.6$, and since the constraint sensitivity varies linearly with time, $t = 1$ is chosen. Figure 6.6 presents the absolute values of the sensitivities (S_g and S_r) over the mesh generated to represent the agent's position. The white circular area in the middle represents the infeasible region. It can be observed that RCS (S_r) is activated when the agent gets closer to the obstacle, as opposed to the constraint sensitivity, which only captures the sensitivity in the constraint value. Furthermore, it has also been observed (though not presented here for the sake of brevity) that by just penalizing the constraint sensitivity using the cost in Equation 6.32, conservative trajectories cannot be obtained



(a) Trajectories



(b) RCS - $S_r^2(t)$



(c) Constraint sensitivity - $S_g^2(t)$

Figure 6.7: Results for constraint desensitized 2D path planning

as the sensitivity profile over-constrains the problem (see Figure 6.6(a)). In the case of penalizing RCS, the agent has sufficient incentive to move away from the obstacle as the variations closer to the constraint boundary now incur a higher penalty (see Figure 6.6(b)).

Figure 6.7 shows the results obtained for the time-optimal path planning problem with $(a_0, b_0) = (0, 0)$, $(a_f, b_f) = (0, 10)$, $c = 5$, $d = 2$, and $\mathrm{v} = 1$. The nominal value of the uncertain parameter $\hat{\rho}$ is chosen to be 0.25. The cost function in Equation 6.36 is minimized with $\phi(x(t_f), t_f) = t_f$, $L(x(t), u(t), t) = 0$, given the dynamics in Equation 6.37, and the constraint in Equation 6.39. The optimal control package GPOPS-II [170] is used to numerically solve the optimal control problem. From Figure 6.7(a), it can be observed

that as the weighting factor Q for the RCS term in the cost function increases, the optimal trajectory becomes more conservative (the distance to the obstacle is greater), and the magnitude of the RCS reduces along the trajectory (see Figure 6.7(b)). Figure 6.7(c) suggests that the conservative trajectories have higher constraint sensitivity, and further corroborates the underlying intuition behind introducing the relevance function.

For this particular example, five different relevance functions: 1) e^{-w^2} (Gaussian); 2) $\max(0, 1 - |w|)$ (Hat function); 3) $s(w)$ (Logistic function); 4) $1/(1 + w^2)$; and 5) $1/(1 + |w|)^2$; were also evaluated to observe the behavior of the constraint desensitized trajectories. All the aforementioned functions provide conservative trajectories similar to the ones in Figure 6.7(a), with slight differences in curvature. The results are not presented in the interest of brevity. Analyzing different relevance functions and their impact on constraint desensitization in general optimal control problems is a separate study meant for future work.

6.4.2 Dependency on the Constraint Form

In this section, the effect of the form of the constraint function over the behavior of the optimal trajectories obtained from constraint desensitized planning is investigated. To this end, the constraint function in Equation 6.39 is expressed alternatively as

$$g_\beta = r_o^\beta - \left[(x-c)^2 + (y-z)^2\right]^{\beta/2} \leq 0, \tag{6.41}$$

where $\beta > 0$. We first analyze the RCS plots shown in Figure 6.8(a), Figure 6.8(c), and Figure 6.8(e). The simulation parameters remain the same as the ones used for the results in Figure 6.6. The general expression for RCS, for any $\beta > 0$, is given by

$$S_r = - \underbrace{r(g_\beta)}_{\text{relevance term}} \times \underbrace{\beta(y-z)t\left[(x-c)^2 + (y-z)^2\right]^{\beta/2-1}}_{\text{constraint sensitivity}} \tag{6.42}$$

Note that, for $\beta > 0$, the relevance term (see Equation 6.42) decays exponentially as the agent moves away from the obstacle. For $\beta > 2$, the constraint sensitivity term increases super-linearly with separation between the agent and the obstacle. Consequently, RCS decays with Euclidean distance and becomes prominent as the agent gets closer to the obstacle. Also, as β increases, the decay (logistic) term overpowers the constraint sensitivity term and the penalty region around the obstacle shrinks.

The optimal trajectories given the constraint Equation 6.41, for $\beta = 0.5$, 2, 4, and with different weights (Q) in the RCS cost can be seen in Figure 6.8(b), Figure 6.8(d), and Figure 6.8(f), respectively. The simulation parameters are the same as the ones used for the results in Figure 6.7. Similar to the results obtained for $\beta = 1$, the optimal trajectories become conservative for all values of β, as Q increases. However, it is noted that the behavior of these optimal trajectories vary. As β increases (for $\beta \geq 1$), the curvature of the desensitized trajectories reduces. For this particular example, under turn radius constraints, the designer can alternatively tune the value of β to obtain the desired trajectory shape.

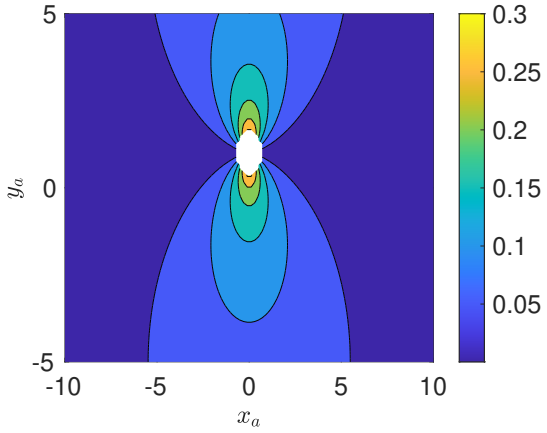
6.4.3 The Car vs. Train Problem

Consider the 1D version of the problem described in subsection 6.4.1, where an agent (car) is restricted to move along the x-axis, and the obstacle (train) is moving along the y-axis. Note that the dynamics for the obstacle remain the same, given in Equation 6.38, while the agent dynamics takes the form

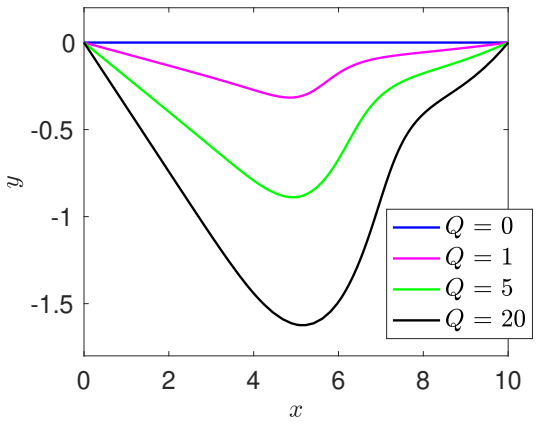
$$\dot{x}(t) = u(t), \quad x_a(0) = a_0,$$

where $u(t) \in [0, \mathrm{v}_{\max}]$, and $y(t) = 0$, $t \in [0, t_f]$. In this case, the state vector $[x, z]^\top$ is two-dimensional.

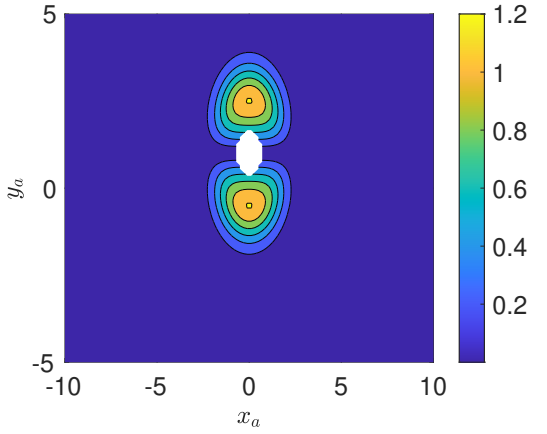
Similar to the previous experiments, we assume that the agent's primary task is to minimize travel time whilst reducing risk of collision under uncertainty in the obstacle's speed.



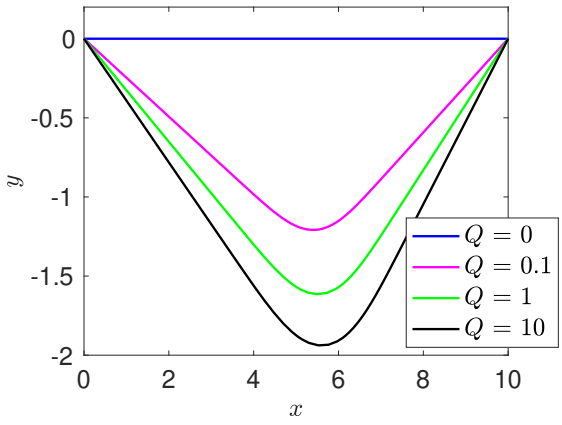
(a) $\beta = 0.5$



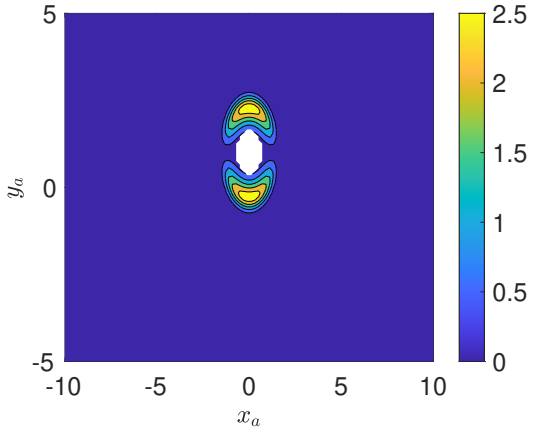
(b) $\beta = 0.5$



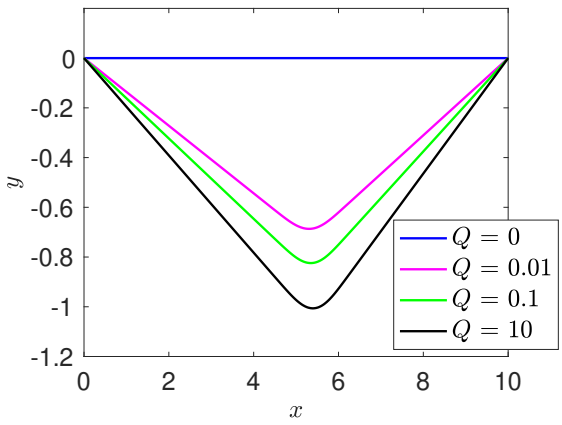
(c) $\beta = 2$



(d) $\beta = 2$

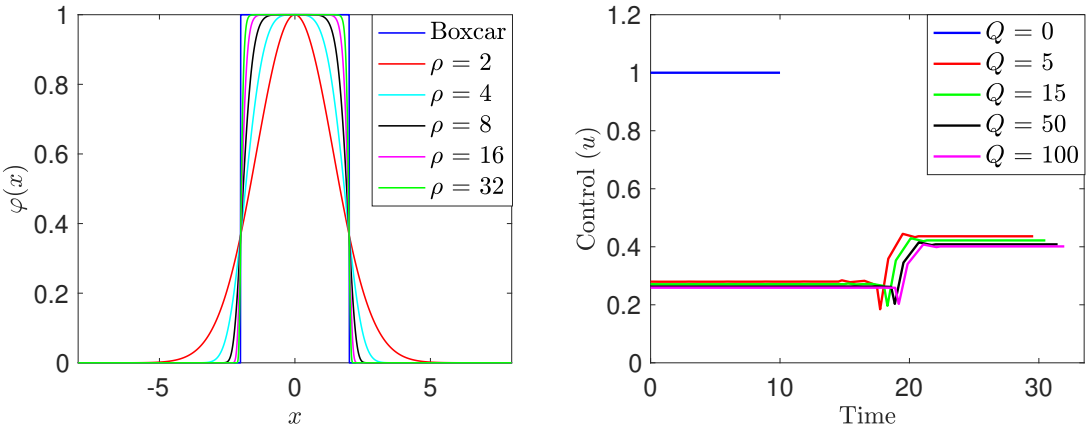


(e) $\beta = 4$



(f) $\beta = 4$

Figure 6.8: Absolute values of the RCS (a, c, e) and the optimal trajectories (b, d, f) for different constraint forms



(a) A super-Gaussian approximation to a Boxcar function

(b) Optimal control for different levels of constraint desensitization

Figure 6.9: The car vs. train problem

It is intuitive to expect that the desensitized solution will ensure that the distance between the train and the car is sufficiently large during the event of crossing the rail track. Computing an RCS cost using the constraint (Equation 6.39) is found to provide a desensitized solution that drives the agent to reach the target point in minimum time, regardless of the distance between the agent and the obstacle (i.e, beyond the safe distance r_o).

The discrepancy between the desensitized solution for the RCS, obtained using the constraint in Equation 6.39 and the intuitive solution can be understood by considering the behavior of a real-world driver. The expression in Equation 6.39, although a valid constraint, does not capture a driver's perception of the collision constraint in this problem. Effectively, the train's motion along the rail track is of no consequence to the driver, except when he is crossing the track. It is during this crossing phase that the driver would ensure sufficient separation (at least the safe distance r_o) between the train and the car to prevent collision.

This motivates us to propose a constraint of the form,

$$\mathbb{1}[|x(t) - c| \leq w_o] \underbrace{[r_o^2 - (y - z(t))^2]}_{g(z)} \leq 0, \tag{6.43}$$

where $\mathbb{1}[\cdot]$ is the indicator function, and $2w_o$ is the width of the track. Note that $\mathbb{1}[|x(t) - c| \leq w_o]$ is a boxcar function which makes the constraint function in Equation 6.43 non-smooth. To this end, we suggest to use the so-called *super-Gaussian* [187]

$$\varphi(x(t); c, w_o) = \exp\left(-\left[\frac{x(t) - c}{w_o}\right]^{\gamma}\right), \tag{6.44}$$

where $\gamma \in 2\mathbb{Z}^+$ (set of positive even numbers), as an approximation to the boxcar function. From Figure 6.9(a), it can be observed that as $\gamma \to \infty$, φ converges to a boxcar function. Subsequently, the RCS for the constraint function in Equation 6.43 with a super-Gaussian approximation can be obtained as

$$S_r = r\big(\varphi(x(t); c, w_o)g(z)\big)\frac{\partial\big(\varphi(x(t); c, w_o)g(z)\big)}{\partial\rho}. \tag{6.45}$$

Figure 6.9(b) shows the optimal control for different levels of constraint desensitization, while employing RCS in Equation 6.45 with the same simulation parameters as before, except now $\gamma = 20$, $\mathrm{v}_{\max} = 1$. It is observed that when there is no penalty on RCS, the car is dangerously close to the train at the crossing. The result further confirms that the RCS in Equation 6.45, when penalized appropriately, allows the car to maintain a safe distance while crossing the track to avoid collision under uncertainty in the speed of the train. This example indicates that an appropriate constraint function is crucial for the success of the proposed approach.

6.4.4 Trade-off Studies with Multiple Obstacles

In this section, the proposed approach is evaluated in instances involving multiple dynamic obstacles moving with uncertain velocities. The dynamics of the agent follow the ODEs in Equation 6.37. In addition to the agent's heading θ, its speed $\mathrm{v} \in [0, 1]$ is included as a control input. Starting at location $(0, 0)$, the agent is tasked with reaching the target location $(30, 0)$ in minimum time while avoiding the obstacles. We consider four different

instances with the number of obstacles $N \in \{2, 3, 5, 10\}$. The obstacles are all assumed to be identical and their movement is restricted to be parallel to the y-axis with their speed ρ being the uncertain parameter. The nominal value of $\hat{\rho}$ is 0.25. A schematic of the environment, containing the initial positions of the agent and the obstacles, and the directions of the obstacles' nominal velocity vectors for the case of $N = 10$, is shown in Figure 6.10.

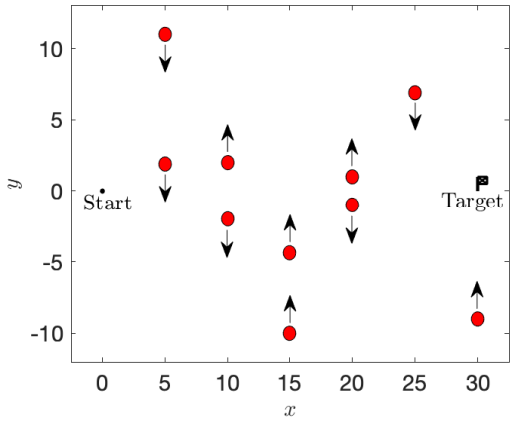
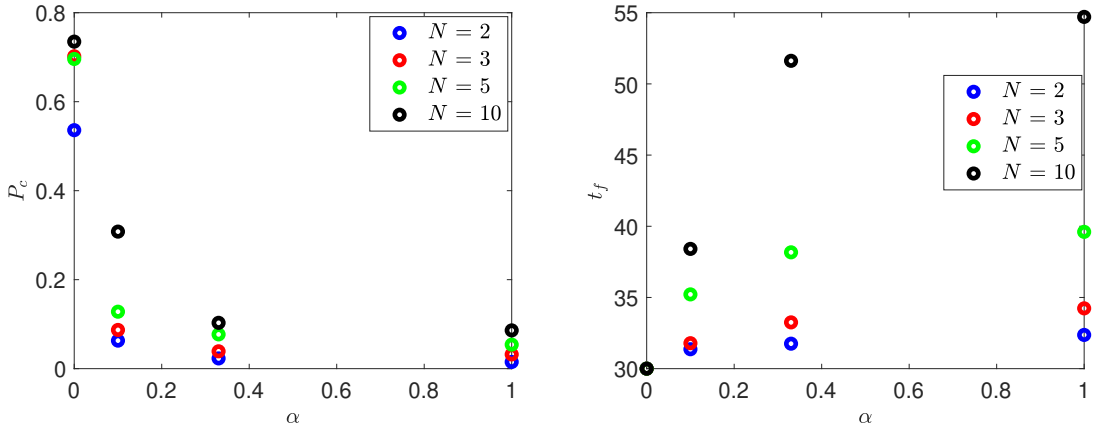


Figure 6.10: Schematic of an uncertain multi-obstacle environment

Note that at each instance, there are N constraints enforcing collision avoidance, and the speed of the associated obstacle is the uncertain parameter. Ignoring the zeros in the matrix $S_r(t)$, all the sensitivity terms are weighted equally by choosing the Q in the RCS cost in Equation 6.36 to be of form $Q = \alpha I_N$. Each instance involves four levels of penalization of the RCS cost with $\alpha = 0$ (blue, no penalty), 0.1 (magenta), 0.33 (green), 1 (black). From Figure 6.11(b), and for different instances ($N = 2, 3, 5, 10$), it can be observed that as α increases, the agent takes longer paths, essentially trying to avoid obstacles while maintaining some safety buffer. To characterize safety, collision probabilities were computed by running Monte Carlo simulations on the optimal trajectory, obtained from GPOPS-II, while propagating the dynamics in an open-loop fashion for 1000 samples. In the Monte Carlo simulations the variation in each of the obstacle's speed $\Delta\rho$ is obtained by sampling from a normal distribution $N(0, \sigma^2)$ with $\sigma^2 = 0.1$. The trade-off between travel times (t_f) and collision probabilities (P_c) for the four instances are shown in Figure 6.11.



(a) Collision probability (P_c)

(b) Travel time (t_f)

Figure 6.11: Trade-offs plots for instances with different number of obstacles

From Figure 6.11, it is seen that penalizing RCS (for higher α values) yields safer trajectories, which mitigate the chance of constraint violation in uncertain environments while trading off optimality (travel time). For instances with $N = 2, 3$, a 95% reduction in collision probability is achieved for a 10% trade-off in travel time. Due to the cumulative effect of increasing the number of obstacles obstacles and the number of uncertain parameters on the RCS cost (that measures the risk of constraint violation), the travel times for safer trajectories are seen to increase with N.

It is observed that the computation times for the tested instances are of the same order of magnitude (a few milliseconds). The approach is limited by the efficiency of the chosen optimal control solver. The regularizer has no guarantees in terms of convexity, and consequently the optimizer may converge to a local minimum. Depending on the initialization, the homotopy class of the obtained trajectories may vary. For the above simulations, we report the optimal trajectory among the ones obtained from different initializations. While in the above simulations the obstacles are restricted to follow simple paths parallel to the y axis, it is important to note that the regularizer can be derived for arbitrary obstacle motion as long as its dynamics are known and the uncertain parameters are identified.

6.5 C-DOC: Co-state Desensitized Optimal Control

In this section, we are interested in desensitizing the cost itself. By denoting

$$\mathcal{J}(u,\rho) = \int_{t_0}^{t} L(x(s),u(s),s)\,\mathrm{d}t + C(\tilde{x}(t),u,t), \tag{6.46}$$

$$C(\tilde{x}(t),u,t) = \int_{t}^{t_f} L(x(s),u(s),s)\,\mathrm{d}t + \phi(x(t_f),t_f), \tag{6.47}$$

we immediately notice that the parametric variation at time t, affects the total cost $\mathcal{J}(u,\rho)$ only through the cost-to-go $C(\tilde{x}(t),u,t)$. Thus, the sensitivity of the total cost for a parametric variation at time t from its nominal value $\hat{\rho}$ can be captured through the term

$$S_C(x(t),\hat{\rho},u,t) = \frac{\partial C}{\partial \rho}(\tilde{x}(t),u,t)\Big|_{\rho=\hat{\rho}}. \tag{6.48}$$

There are several ways to capture the effect of the parametric variations on the cost, one of which is to consider the following sensitivity cost

$$\mathcal{J}_c(u,\hat{\rho}) = \int_{t_0}^{t_f} \|S_C(x(t),\hat{\rho},u,t)\|^2_{Q(t)}\,\mathrm{d}t, \tag{6.49}$$

for some $Q(t) \geq 0$, for all $t \geq t_0$.

There are three major formulations relevant to the problem of cost-based desensitization, which are as follows.

Problem 3. *Solve*

$$\inf_{u\in\mathcal{U}} \;\; \mathcal{J}(u,\hat{\rho}) \tag{6.50a}$$

$$\textit{subject to} \;\; \mathcal{J}_c(u,\hat{\rho}) \leq D. \tag{6.50b}$$

Let us denote the solution of Problem 3 to be the "cost-desensitization" function $\mathsf{J}(D)$ which represents the optimal cost given a bound on the sensitivity metric. A similar prob-

lem is to consider minimizing the sensitivity of the cost for a given bound on the performance index, as presented below.

Problem 4. *Solve*

$$\inf_{u \in \mathcal{U}} \ \mathcal{J}_c(u, \hat{\rho}) \tag{6.51a}$$

$$\textit{subject to} \ \ \mathcal{J}(u, \hat{\rho}) \leq J. \tag{6.51b}$$

Let us denote the solution of Problem 4 to be the "desensitization-cost" function $\mathsf{D}(J)$. Finding analytical or numerical solutions to $\mathsf{J}(D)$ or $\mathsf{D}(J)$ are challenging. However, $\mathsf{J}(D)$ or $\mathsf{D}(J)$ can be constructed by solving the following family optimization problems for all $\alpha \in [0, \infty)$.

Problem 5. *Solve*

$$\inf_{u \in \mathcal{U}} \mathcal{J}(u, \hat{\rho}) + \alpha \mathcal{J}_c(u, \hat{\rho}) \tag{6.52}$$

By observing that the scalar α can be absorbed into the matrix $Q(t)$, we will rewrite the objective function in Problem 5 as

$$\mathcal{J}_s(u) = \mathcal{J}(u, \hat{\rho}) + \mathcal{J}_c(u, \hat{\rho}).$$

When the sensitivity cost has zero weight ($Q(t) \equiv 0$), we solve problem (Equation 6.1) and retrieve $\limsup_{D \to \infty} \mathsf{J}(D)$, and as we increase the weight on the sensitivity cost (through $Q(t)$), we arrive at an optimal control whose performance is more insensitive to the variations in the parameters. In the limit when $Q(t) \to \infty$ for all t, we retrieve $\limsup_{J \to \infty} \mathsf{D}(J)$. In this work, we will focus on minimizing $\mathcal{J}_s(u)$. Detailed analysis of $\mathsf{J}(D)$ and $\mathsf{D}(J)$ will appear elsewhere.

The new optimization problem we are interested in solving is

$$\inf_{u} \quad \mathcal{J}_s(u) \tag{6.53a}$$

$$\text{subject to } \dot{x} = f(x, \hat{\rho}, u, t), \quad x(t_0) = x_0 \tag{6.53b}$$

$$\psi(x(t_f), t_f) = 0. \tag{6.53c}$$

The following section presents a formal theorem for the fact that the co-states capture the sensitivity of the cost-to-go function for any given control input $\bar{u}(t)$, that satisfies the terminal constraint in Equation 6.53c with nominal value of the parameter $\hat{\rho}$. In this section, the trajectory constraints in Equation 6.2b are ignored. The result would allow us to penalize a weighted norm of the co-states, with their dynamics obtained from the adjoint equations, that desensitizes the cost function with respect to the variations in the targeted parameters.

6.5.1 Main Theorem

In this section we characterize the cost-sensitivity $S_C(x(t), \hat{\rho}, u, t)$ in terms of the co-state process associated with the optimal control problem given by Equation 6.1, Equation 6.2a, and Equation 6.2c. The following theorem shows that the sensitivity of the cost-to-go function with respect to the state at time t can be represented by a co-state process λ with certain boundary conditions at the final time t_f.

Theorem 7. *Consider the dynamical system $\dot{x} = f(x, u, t)$, evolving under a given control law $\bar{u} \in \bar{\mathcal{U}} \subseteq \mathcal{U}$, where*

$$\bar{\mathcal{U}} = \Big\{ \bar{u} : [t_0, t_f] \to \mathbb{R}^m \textit{ is PWC}, \bar{u}(t) \in U, \ \psi(x(t_f), t_f) = 0, \\ x(t_f) = x_0 + \int_{t_0}^{t_f} f(x(t), \bar{u}(t), t)\, \mathrm{d}t \Big\}.$$

Then, for a cost-to-go function (associated with the cost functional in Equation 6.1) with

$x = x(t)$

$$C(x, \bar{u}, t) = \phi(x(t_f), t_f) + \int_t^{t_f} L(x(\tau), \bar{u}(\tau), \tau)\, \mathrm{d}\tau, \tag{6.54}$$

under the control $\bar{u} \in \bar{\mathcal{U}}$, the sensitivity of the cost-to-go function with respect to the state x at time t is,

$$\lambda^\top(t) = \frac{\partial C}{\partial x}(x(t), \bar{u}, t), \tag{6.55}$$

which obeys the dynamics

$$\dot{\lambda}^\top(t) = -\frac{\partial H}{\partial x}(x(t), \bar{u}, \lambda(t), t), \tag{6.56}$$

where

$$H(x, u, \lambda, t) = L(x, u, t) + \lambda^\top f(x, u, t). \tag{6.57}$$

Furthermore, the terminal condition for the ODE in Equation 6.56 is given by

$$\lambda(t_f) = \frac{\partial \phi}{\partial x}(x(t_f), t_f). \tag{6.58}$$

Proof. For a fix control $\bar{u}$, the cost-to-go from any state x at time t is

$$C(x, \bar{u}, t) = \phi(x(t_f), t_f) + \int_t^{t_f} L(x(s), \bar{u}(s), s)\mathrm{d}s$$

where $x(t) = x$.

Let the perturbed state at time t be represented by $x(t, \alpha) = x(t) + \alpha\delta x(t)$ where $\alpha \in [0, \alpha_0)$ for some $\alpha_0 > 0$, and $\delta x(t) \in \mathbb{R}^n$. With this perturbation the new cost-to-go is

$$C(x + \alpha\delta x, \bar{u}, t) = \phi(x(t_f, \alpha), t_f) + \int_t^{t_f} L(x(s, \alpha), \bar{u}(s), s)\mathrm{d}s$$

where $x(s,\alpha)$ denotes the perturbed state at time $s \geq t$. By denoting $x(s,\alpha) = x(s,0) + \alpha\delta x(s)$, for all $s \geq t$, we obtain

$$\delta\dot{x}(s) = f_x(x(s,0),\bar{u}(s),s)\delta x(s) + O(\alpha),$$

where $O(\alpha)$ is such that $\lim_{\alpha\to 0} O(\alpha) = 0$. Consequently, $\delta x(s) = \Gamma(s,t)\delta x(t) + O(\alpha)$ where $\Gamma(s,t)$ is the state transition matrix corresponding to the matrix $f_x(x(s,0),u(s),s)$.

Therefore,

$$\begin{aligned} C(x+\alpha\delta x,\bar{u},t) - C(x,\bar{u},t) &= \alpha\phi_x(x(t_f,0),t_f)\Gamma(t_f,t)\delta x \\ &+ \alpha\left[\int_t^{t_f} L_x(x(s,0),u(s),s)\Gamma(s,t)\mathrm{d}s\right]\delta x + O(\alpha^2) \end{aligned}$$

and thus,

$$\begin{aligned} \lim_{\alpha\to 0^+} \frac{C(x+\alpha\delta x,\bar{u},t) - C(x,\bar{u},t)}{\alpha} &= \Big[\phi_x(x(t_f,0),t_f)\Gamma(t_f,t) \\ &+ \int_t^{t_f} L_x(x(s,0),u(s),s)\Gamma(s,t)\mathrm{d}s\Big]\delta x, \end{aligned}$$

and

$$C_x(x,\bar{u},t) = \phi_x(x(t_f,0),t_f)\Gamma(t_f,t) + \int_t^{t_f} L_x(x(s,0),\bar{u}(s),s)\Gamma(s,t)\mathrm{d}s.$$

At this point, if we denote

$$\lambda^\top \triangleq C_x(x,\bar{u},t).$$

We then have

$$\dot{\lambda}^\top = -\lambda^\top f_x - L_x,$$

since $\dot{\Gamma}(s,t) = -\Gamma(s,t)f_x(x(t,0),u(t),t)$. Furthermore, $\lambda(s)$ satisfies the terminal condi-

tion $\lambda(t_f) = \phi_x(x(t_f), t_f)$. Thus, if we define the Hamiltonian as $H = L + \lambda^\mathsf{T} f$, it follows that

$$\dot{\lambda}^\top = -\frac{\partial H}{\partial x}$$

and this λ represents the first order variation in cost-to-go with boundary condition

$$\lambda(t_f) = \phi_x(x(t_f), t_f).$$

The result follows. □

It is interesting to note that the theorem holds not only for the optimal control (a result that follows directly from the maximum principle [188]), but for any control law that is piecewise continuous and ensures that the terminal constraint is met. The C-DOC problem can now be fully formulated using this result.

For the C-DOC problem the augmented state is $\tilde{x} = [x^\top \ \ \rho^\top]$ with dynamics given in Equation 6.4. The Hamiltonian, defined in Theorem 7, for this system, can be written as

$$\begin{aligned} H(\tilde{x}, u, \lambda, \mu, t) &= L(x, u, t) + \lambda^\top \dot{x} + \mu^\top \dot{\rho} \\ &= L(x, u, t) + \lambda^\top f(x, \rho, u, t), \end{aligned} \tag{6.59}$$

where λ and μ are the co-states corresponding to state x and vector of parameters defined by ρ, respectively. The corresponding adjoint equations are given by

$$\dot{\lambda}^\top = -\frac{\partial H}{\partial x}(\tilde{x}, u, \lambda, \mu, t) = -\lambda^\top \frac{\partial f}{\partial x}(x, \rho, u, t) - \frac{\partial L}{\partial x}(x, u, t), \tag{6.60a}$$

$$\dot{\mu}^\top = -\frac{\partial H}{\partial \rho}(\tilde{x}, u, \lambda, \mu, t) = -\lambda^\top \frac{\partial f}{\partial \rho}(x, \rho, u, t). \tag{6.60b}$$

Since the co-states represent the sensitivity of the cost-to-go function for a given control

input $u(t)$ (Theorem 7), they can be expressed as

$$\lambda(t)^\top = \frac{\partial C}{\partial x}(\tilde{x}(t), u, t), \tag{6.61a}$$

$$\mu(t)^\top = \frac{\partial C}{\partial \rho}(\tilde{x}(t), u, t), \tag{6.61b}$$

for a given control $u \in \bar{\mathcal{U}}$, this results in the trajectory $x(t)$ for $t_0 \leq t \leq t_f$, where

$$C(\tilde{x}, u, t) = \phi(x(t_f), t_f) + \int_t^{t_f} L(x(\tau), u(\tau), \tau)\, \mathrm{d}\tau.$$

Note that ρ is an augmented state in the given problem and affects the cost $\mathcal{J}$ through the state x, whose dynamics is a function of ρ. Since we have used $\dot{\rho} = 0$ and $\rho(t_0) = \hat{\rho}$, we have ensured that $\rho(t) = \hat{\rho}$. Thus, by comparing Equation 6.48 and Equation 6.61b, we obtain $\mu(t) = S_C(x(t), \hat{\rho}, u, t)$. Therefore, weighting the co-state in the existing cost function will ensure that the solution of the augmented problem minimizes the sensitivity of the cost $\mathcal{J}$ with respect to parametric variations. This results in an updated optimal control problem with an augmented cost, accounting for the sensitivity component, given by

$$\mathcal{J}_s(u) = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} \Big[L(x(t), u(t), t) + \mu^\top(t) Q(t) \mu(t) \Big]\, \mathrm{d}t. \tag{6.62}$$

Minimizing the cost in Equation 6.62, subject to the dynamics in Equation 6.4, terminal constraint in Equation 6.2c, and the transversality conditions in Equation 6.58 with

$$\mu(t_f) = 0, \tag{6.63}$$

yields a desensitized optimal control problem for the original problem. Here, $Q(t) \in \mathbb{R}^{\ell \times \ell}$

is a user-defined positive semi-definite weighting function and is generally of the form

$$Q(t) \equiv \text{diag}(\alpha_1(t), \ldots, \alpha_\ell(t)). \tag{6.64}$$

This co-state based approach requires formulating $2(n+\ell)$ number of states, as compared to the higher $2(n+\ell)^2+n+\ell$ states in [111], employing sensitivity matrices for an optimal control problem. The resulting problem is typically solved by the off-the-shelf existing solvers.

6.5.2 Numerical Examples

The following section presents some numerical examples that will aid in understanding the implementation of this technique and will elucidate its subtleties. The simulations are obtained using GPOPS-II [170].

Consider an optimal control problem of minimizing a quadratic cost

$$\mathcal{J}(u) = \int_0^{t_f} \frac{1}{2}(x^\top R_1 x + u^\top R_2 u)\,\mathrm{d}t, \tag{6.65}$$

given the n-dimensional linear dynamics with parameter vector ρ

$$\dot{x} = A(\rho)x + B(\rho)u, \tag{6.66a}$$

$$\dot{\rho} = 0, \tag{6.66b}$$

with initial conditions $x(0) = x_0, \quad \rho(0) = \hat{\rho}$, where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $\rho \in \mathbb{R}^\ell$, $A : \mathbb{R}^\ell \to \mathbb{R}^{n\times n}$, $B : \mathbb{R}^\ell \to \mathbb{R}^{n\times m}$, $R_1 \geq 0$, $R_2 > 0$, and t_f is fixed. The goal is to desensitize the cost with respect to the parameter ρ. Following the steps to construct the cost term for

desensitization, the Hamiltonian is given by

$$\begin{aligned} H &= \frac{1}{2}(x^\top R_1 x + u^\top R_2 u) + \lambda^\top \dot{x} + \mu^\top \dot{\rho}, \\ &= \frac{1}{2}(x^\top R_1 x + u^\top R_2 u) + \lambda^\top (A(\rho)x + B(\rho)u). \end{aligned} \tag{6.67}$$

The adjoint equations are

$$\dot{\lambda}^\top = -\frac{\partial H}{\partial x} = -x^\top R_1 - \lambda^\top A(\rho), \tag{6.68}$$

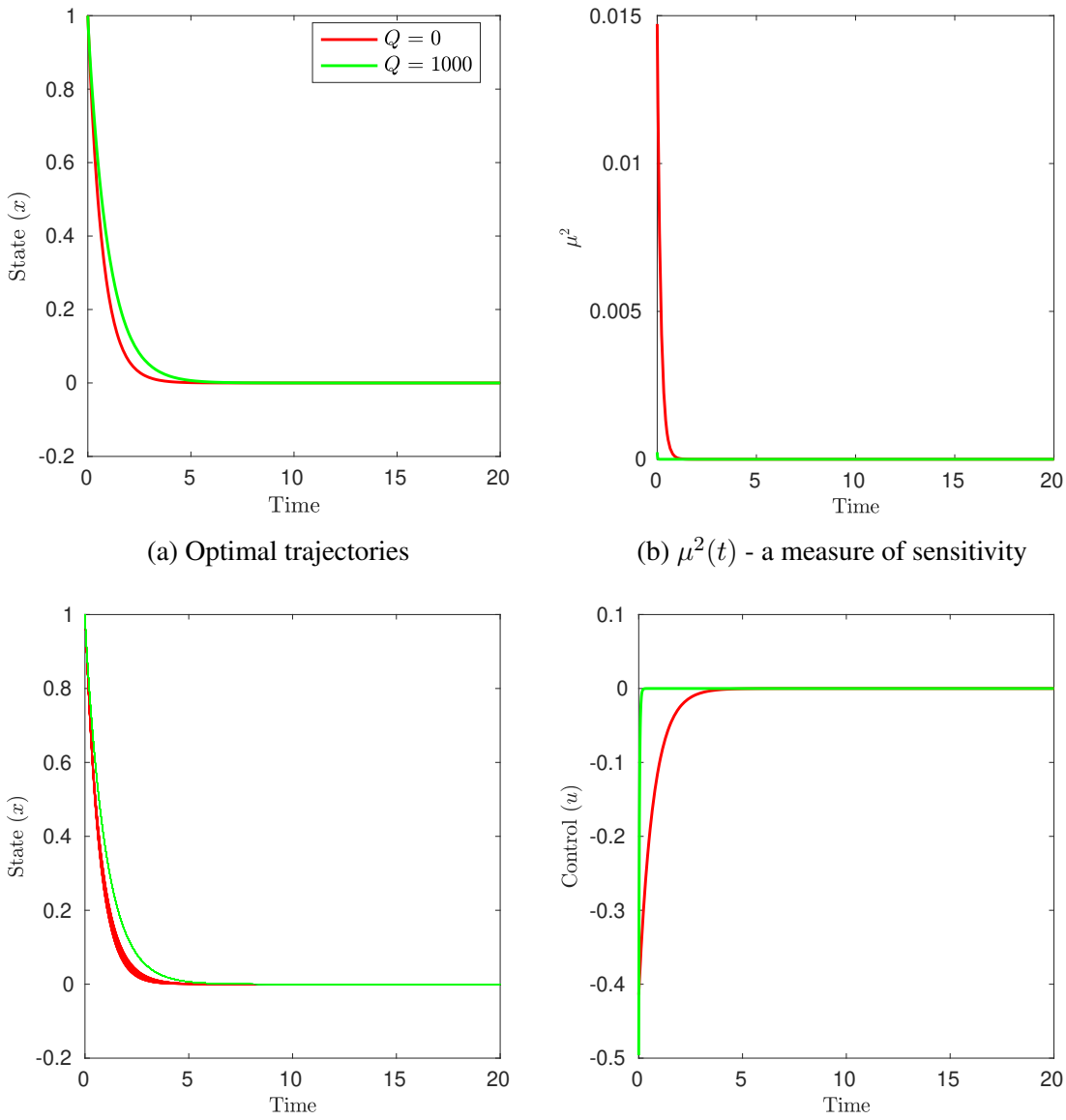
$$\dot{\mu}^\top = -\frac{\partial H}{\partial \rho} = -(x^\top \otimes \lambda^\top)\frac{\partial}{\partial \rho}\text{vec } A(\rho) - (u^\top \otimes \lambda^\top)\frac{\partial}{\partial \rho}\text{vec } B(\rho). \tag{6.69}$$

where λ and μ are the co-states of x and ρ, respectively. Since the cost has to be desensitized with respect to ρ, the augmented cost that has to be minimized for the C-DOC problem is given by

$$\mathcal{J}_s(u) = \int_0^{t_f} \frac{1}{2}(x^\top R_1 x + u^\top R_2 u + \mu^\top Q \mu)\, \mathrm{d}t. \tag{6.70}$$

To demonstrate the results, we consider a one-dimensional linear system with the dynamics $\dot{x} = ax + bu$ with initial condition $x(0) = 1$, and let $R_1 = R_2 = 2$, $t_f = 20$. We first analyze the case where b is the uncertain parameter with its nominal value as $b_0 = 1$, and $a = -1$. The solutions obtained for $Q = 0$ and $1,000$ are shown in Figure 6.12. Note that the sensitivity measure ($\mu^2(t)$) in Figure 6.12(b) is lower for the desensitized solution. Since b is the source of uncertainty that perturbs the trajectory (and eventually the cost), by introducing desensitization ($Q = 1,000$), it can be observed from Figure 6.12(d) that the control goes to zero earlier compared to the non-desensitized solution. By making the control zero, the source of uncertainty is removed from the system. The results obtained from the Monte-Carlo simulations with $b \in [0.8b_0, 1.2b_0]$ are shown in Figure 6.12(c), which suggests that the variation in the cost for the desensitized solution is significantly lower.

The results for the case where a is the uncertain parameter with its nominal value as



(a) Optimal trajectories

(b) $\mu^2(t)$ - a measure of sensitivity

(c) Monte-Carlo simulations

(d) Optimal Control (u^*)

Figure 6.12: Results obtained for an LQR problem with B matrix being uncertain

$a_0 = -1$ (stable), and $b = 1$ are shown in Figure 6.13. Since a is the source of uncertainty, by switching on the desensitization ($Q = 1,000$), it can be observed from Figure 6.13(a) that the state approaches zero faster compared to the non-desensitized solution. Consequently, from the Monte-Carlo simulations ($a \in [0.8a_0, 1.2a_0]$), it can be observed that the variations in the optimal trajectory (Figure 6.13(c)), and the cost (Figure 6.13(d)) are significantly lower for the desensitized solution, though the cost for the same is higher which is a trade-off. The error bars in Figure 6.13(d) represent the minimum and the maximum costs obtained form the Monte-Carlo results where the corresponding grey bars represent the nominal costs with $a = a_0$.

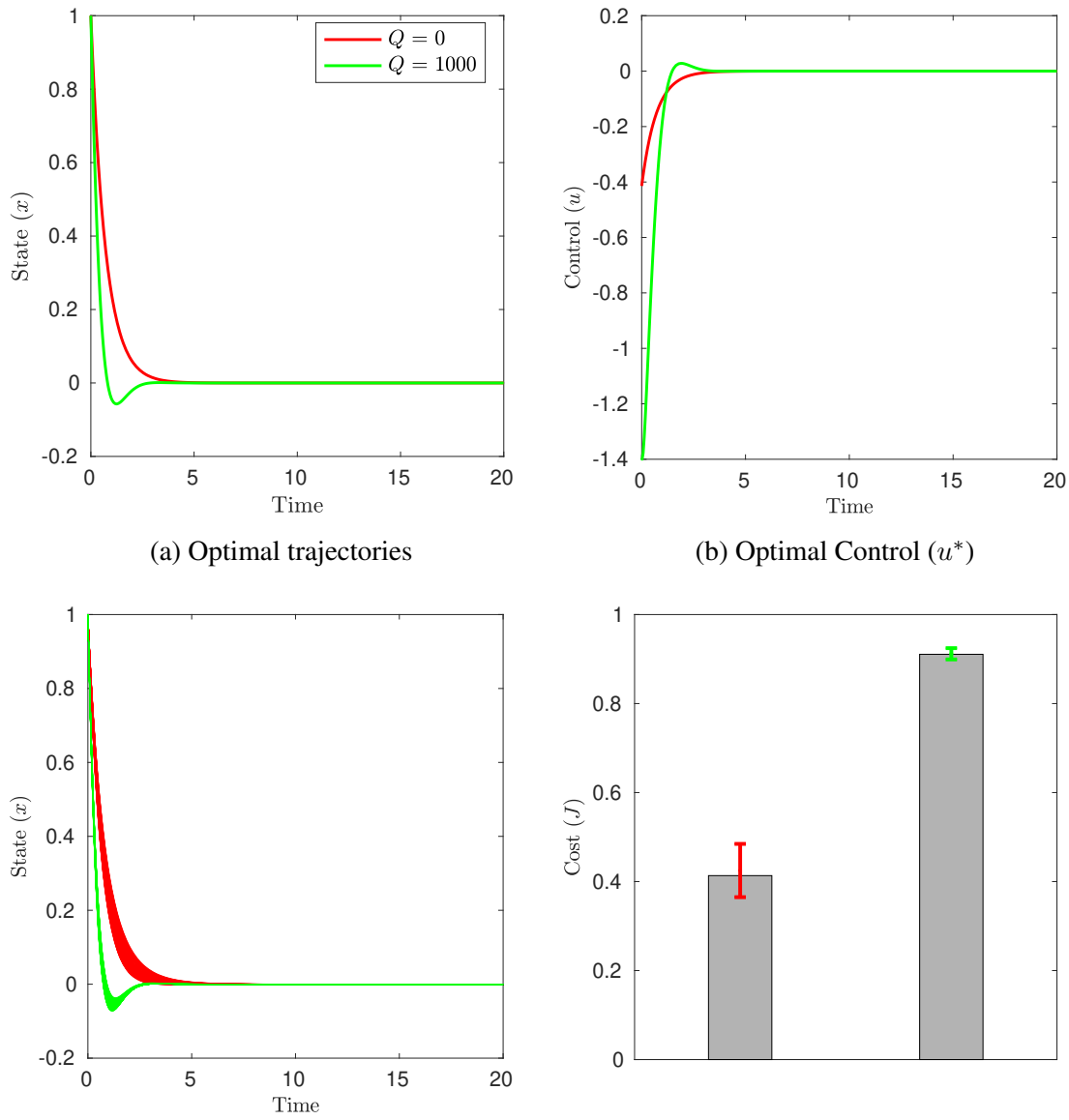
A more interesting case is a marginally stable system with $a_0 = 0$, and $a \in [-0.2, 0.2]$. The corresponding results can be found in Figure 6.14. In the previous cases, although a parametric variation in a is studied, such variations did not change the stability of the system, i.e., if the nominal system is stable, then the system with parametric variation is stable as well. Since a can be both stable and unstable, the optimal control obtained for the nominal system without desensitization will be less effective combating the instabilities compared to the desensitized solution, as can be seen from the dispersion in trajectories (and costs) in Figure 6.14.

6.5.3 Discussion

In this section we address the relationship between the sensitivity matrix defined in Equation 6.6 and the co-states λ. Let us note that, $\lambda^\top(t) = \frac{\partial C}{\partial x}(x(t), u, t)$, which can be expressed as

$$\lambda^\top(t) = \frac{\partial C(x(t), u, t)}{\partial x_0} \left[\frac{\partial x(t)}{\partial x_0}\right]^{-1} = \frac{\partial C(x(t), u, t)}{\partial x_0} \Lambda(t|t_0, x_0)^{-1}, \tag{6.71}$$

where $\Lambda(t|t_0, x_0) = \dfrac{\partial x(t)}{\partial x_0}$ is the sensitivity of the state at time t along the trajectory with respect to variation in its initial condition x_0 [111]. Note that the dependency of $x(t)$,



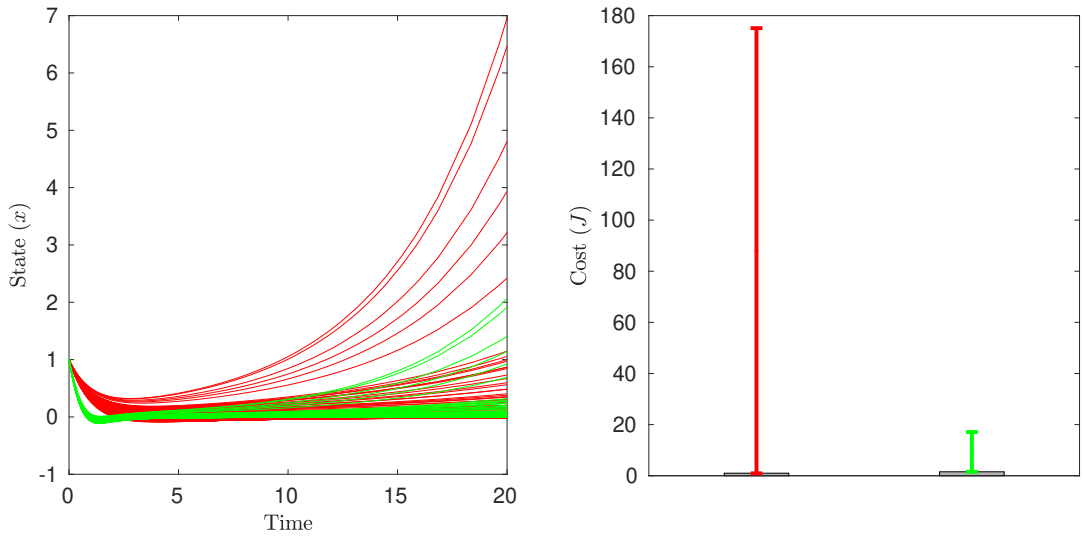
(a) Optimal trajectories

(b) Optimal Control (u^*)

(c) Monte-Carlo simulations

(d) Cost (J)

Figure 6.13: Results obtained for an LQR problem with a stable A matrix being uncertain



(a) Monte-Carlo simulations (b) Cost (J)

Figure 6.14: Results obtained for the LQR problem with $a = 0$ (red: $Q = 0$, green: $Q = 1,000$)

$t \geq t_0$ on t_0 and x_0 (initial conditions) is implicit. The relationship between the co-state and the sensitivity matrix in Equation 6.71 can be generalized to obtain the sensitivity of the solution with regards to the state at any other time t' as

$$\lambda^\top(t) = \frac{\partial C(x(t), u, t)}{\partial x(t')} \frac{\partial x(t')}{\partial x(t)} = \frac{\partial C(x(t), u, t)}{\partial x(t')} \Lambda(t|t', x(t'))^{-1} \; \forall \, t, t' \in [t_0, t_f]. \quad (6.72)$$

Therefore,

$$\lambda(t) = \Lambda(t|t', x(t'))^{-\top} \left[\frac{\partial C(x(t), u, t)}{\partial x(t')} \right]^\top, \quad (6.73)$$

$$\left[\frac{\partial C(x(t), u, t)}{\partial x(t')} \right]^\top = \Lambda(t|t', x(t'))^\top \lambda(t). \quad (6.74)$$

From the above expressions, we observe that the sensitivity matrix $\Lambda(t|t', x(t'))^\top$ is essentially the transition matrix between the co-states $\lambda(t)$ and the partial of the cost-to-go function at time t with respect to the state at time t' i.e., the sensitivity of the cost-to-go function at time t with respect to the state at time $t' < t$.

CHAPTER 7

DESENSITIZED STRATEGIES FOR PURSUIT-EVASION GAMES WITH ASYMMETRIC INFORMATION

7.1 Preliminaries

7.1.1 Differential Games with Asymmetric Information

Consider a two-player differential game with payoff function

$$\mathcal{J}_\rho(u, v) = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} L(x(t), u(t), v(t), t)\,\mathrm{d}t, \tag{7.1}$$

and constraints

$$\dot{x}(t) = f(x(t), \rho, u(t), v(t), t), \quad x(t_0) = x_0, \tag{7.2}$$

$$g^u(x(t), \rho, t) \leq 0, \tag{7.3}$$

$$g^v(x(t), \rho, t) \leq 0, \tag{7.4}$$

where $t \in [t_0,\, t_f]$ denotes time, with t_0 being the fixed initial time and t_f being the final time, $x(t) \in \mathbb{R}^n$ denotes the state of the game, with x_0 being the fixed state at t_0, and $\rho \in \mathcal{D} \subset \mathbb{R}^k$ is a vector of model parameters. The player choosing the control input $u \in \mathcal{U} = \{\text{Piecewise Continuous (PWC)},\, u(t) \in U,\, \forall\, t \in [t_0, t_f]\}$ with $U \subseteq \mathbb{R}^m$, the set of allowable values of $u(t)$, aims to minimize the payoff function in Equation 7.1. On the other hand, the player choosing the control input $v \in \mathcal{V} = \{\text{PWC},\, v(t) \in V,\, \forall\, t \in [t_0, t_f]\}$ with $V \subseteq \mathbb{R}^\ell$, the set of allowable values of $v(t)$, aims to maximize the payoff function in Equation 7.1. In Equation 7.1, $\phi : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ is the terminal cost function, and $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ is the running cost. The subscript ρ indicates that given

the controllers u and v, the payoff in Equation 7.1 is a function of the parameter vector ρ. Finally, $g^u(g^v) : \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^q(\mathbb{R}^r)$, denotes the state constraints that the minimizing (maximizing) player has to meet along the trajectory. It is assumed that the state constraints g^u and g^v depend on the control inputs u and v, respectively, and are independent of the control inputs of the adversary. For the case of pursuit-evasion games in dynamic uncertain environments, which is analyzed in section 7.3, the constraints in Equation 7.3 and Equation 7.4 can be used to enforce collision avoidance of the moving obstacles for both agents. The game ends when the condition

$$\psi(x(t_f), t_f) = 0, \tag{7.5}$$

is met at the final time.

In this work, it is assumed that the minimizing player cannot observe the parameter vector ρ, and it knows only the nominal value of the parameter, which is denoted by $\hat{\rho}$. For the minimizing player, the parameter vector assumes variations about the nominal value $\hat{\rho}$. It is assumed that the maximizing player knows the true value of the parameter, which includes the nominal value ($\hat{\rho}$) and its possible variation $\Delta\rho = \bar{\rho} - \hat{\rho}$, where $\bar{\rho}$ is the true value of the parameter vector. In terms of the information that the players have about their adversaries, in this work we assume that the maximizing player is aware of the fact that the minimizing player only knows the nominal value of the parameter. The minimizing player is aware of the fact that its adversary knows the true value of the parameter vector, although it does not know its true value. Therefore, both players know that the minimizing player is at an information disadvantage.

For the minimizing player, the parametric variations can lead to constraint violations and to this end, it modifies its payoff function by augmenting a sensitivity-function-based risk measure. It is further assumed that the maximizing player is unaware of the modified payoff function that the minimizing player employs. Consequently, in the proposed ap-

proach each player tries to solve a non-zero-sum dynamic game (based on the information each player has about its own game) to obtain the corresponding feedback strategies. In the following subsection, we introduce the theory of sensitivity functions that is appropriate for the differential game formulation.

7.1.2 Sensitivity functions

Consider the dynamics in Equation 7.2, and assume variations in the model parameters $\rho \in \mathcal{D}$, with $\hat{\rho}$ being the nominal value of the parameter vector. Furthermore, assume that $f(x, \rho, u, v, t)$ is continuous in (x, ρ, u, v, t), and continuously differentiable with respect to x and ρ for all $(x, \rho, u, v, t) \in \mathbb{R}^n \times \mathcal{D} \times U \times V \times [t_0, t_f]$. The solution to the differential equation in Equation 7.2 from the initial condition x_0 using admissible controls, $u \in \mathcal{U}$ and $v \in \mathcal{V}$, is given by

$$x(\rho, t) = x_0 + \int_{t_0}^{t} f(x(\rho, \tau), \rho, u(\tau), v(\tau), \tau)\, \mathrm{d}\tau. \tag{7.6}$$

Since $f(x, \rho, u, t)$ is differentiable with respect to ρ, it follows that

$$\begin{aligned} \frac{\partial x}{\partial \rho}(\rho, t) = \int_{t_0}^{t} \Bigg[& \frac{\partial f}{\partial x}(x(\rho, \tau), \rho, u(\tau), v(\tau), \tau) \frac{\partial x}{\partial \rho}(\rho, \tau) \\ & + \frac{\partial f}{\partial \rho}(x(\rho, \tau), \rho, u(\tau), v(\tau), \tau) \Bigg]\, \mathrm{d}\tau. \end{aligned} \tag{7.7}$$

Taking the derivative with respect to t, we obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{\partial x}{\partial \rho}(\rho, t) \right] = & \left. \frac{\partial f(x, \rho, u(t), v(t), t)}{\partial x} \right|_{x = x(\rho, t)} \frac{\partial x}{\partial \rho}(\rho, t) \\ & + \left. \frac{\partial f(x, \rho, u(t), v(t), t)}{\partial \rho} \right|_{x = x(\rho, t)}. \end{aligned} \tag{7.8}$$

Evaluating Equation 7.8 at the nominal conditions ($\rho = \hat{\rho}$), the dynamics for the *parameter sensitivity function* $S : [t_0, t_f] \to \mathbb{R}^{n \times \ell}$

$$S(t) = \left. \frac{\partial x(\rho, t)}{\partial \rho} \right|_{x = x(\hat{\rho}, t)} \tag{7.9}$$

can be obtained as

$$\dot{S}(t) = A(t)S(t) + B(t), \quad S(t_0) = 0_{n \times \ell}, \tag{7.10}$$

where

$$A(t) = \left. \frac{\partial f(x, \rho, u(t), v(t), t)}{\partial x} \right|_{x = x(\hat{\rho}, t),\, \rho = \hat{\rho}}, \tag{7.11}$$

$$B(t) = \left. \frac{\partial f(x, \rho, u(t), v(t), t)}{\partial \rho} \right|_{x = x(\hat{\rho}, t),\, \rho = \hat{\rho}}. \tag{7.12}$$

Since the initial state is given (fixed), the initial condition for the sensitivity function is the zero matrix, and Equation 7.10 is called the *sensitivity equation* in the literature [164]. To compute the sensitivity function over time, the state x has to be propagated using the dynamics in Equation 7.2 under the nominal conditions,

$$\dot{\hat{x}} = f(\hat{x}, \hat{\rho}, u, v, t), \quad \hat{x}(t_0) = x_0. \tag{7.13}$$

Note that the minimizing player u has information only about the nominal value of the parameter. Here, $\hat{x}(t) = x(\hat{\rho}, t)$ denotes the nominal state at time t, as computed by the minimizing player while solving the differential equation in Equation 7.13, given the controllers u and v. From the properties of continuous dependence with respect to the parameters and the differentiability of solutions of ordinary differential equations, for sufficiently

small variations in $\hat{\rho}$, the solution $x(\rho, t)$ can be approximated by

$$x(\rho, t) \approx x(\hat{\rho}, t) + S(t)(\rho - \hat{\rho}). \tag{7.14}$$

This is a first-order approximation of $x(\rho, t)$ about the nominal solution $x(\hat{\rho}, t)$. In the next section we develop a scheme to generate safe trajectories for the minimizing player by penalizing a risk measure that is defined using sensitivity functions.

7.2 Constraint Desensitized Planning

7.2.1 Relevant Constraint Sensitivity

For the differential game expressed using Equation 7.1-Equation 7.5, assuming $g^u(x, \rho, t)$ is a smooth function in x, we obtain the sensitivity of the constraint function as

$$\begin{aligned} S^u(t) &= \left.\frac{\partial g^u(x(\rho,t),\rho,t)}{\partial \rho}\right|_{\rho=\hat{\rho}} \\ &= \left(\frac{\partial g^u(x,\rho,t)}{\partial x}\frac{\partial x(\rho,t)}{\partial \rho} + \frac{\partial g^u(x,\rho,t)}{\partial \rho}\right)\Bigg|_{x=x(\hat{\rho},t),\rho=\hat{\rho}} \\ &= \left(\frac{\partial g^u(x,\rho,t)}{\partial x}S(t) + \frac{\partial g^u(x,\rho,t)}{\partial \rho}\right)\Bigg|_{x=x(\hat{\rho},t),\rho=\hat{\rho}}. \end{aligned} \tag{7.15}$$

In chapter 6, it has been argued that "variations in the constraint value when the system is far from the constraint boundary are not as important as when the system is close to the constraint boundary". Therefore, to account for the fact that the constraint variations are more likely to cause constraint violations when the system is closer to the constraint boundary, a *relevance function* $r : \mathbb{R} \to [0, \infty)$ of the form

$$r(z) = \begin{cases} \tilde{r}(z), & \text{if } z \leq 0, \\ \tilde{r}(0), & \text{if } z > 0, \end{cases} \tag{7.16}$$

where $\tilde{r} : \mathbb{R} \to [0, \infty)$ is a continuous function that is monotonically increasing over the interval $(-\infty, 0]$, that is, $\tilde{r}(z) \geq \tilde{r}(y)$, if $z > y$ for all $z, y \leq 0$, is considered.

Next, we construct the *relevant constraint sensitivity* (RCS) matrix $S^r : [0, \infty) \to \mathbb{R}^{q \times k}$ as

$$S^r(t) = RS^u(t), \tag{7.17}$$

where

$$R = \text{diag}\left(r(g_1^u(x(\rho, t), \rho, t)), \ldots, r(g_q^u(x(\rho, t), \rho, t))\right). \tag{7.18}$$

For the purpose of the analysis, as considered in chapter 6, the derivative of the logistic function $s(z) = 1/(1 + e^{-z})$ is chosen as the candidate relevance function, that is,

$$\tilde{r}(z) = s(z)(1 - s(z)). \tag{7.19}$$

The sensitivity matrix S^r captures the idea of giving more importance to variations near the constraint boundary. Using the RCS, a conservative strategy for the minimizing player is proposed in the following subsection.

7.2.2 Players' Strategies

As mentioned earlier, each player solves a non-zero-sum game based on the information it has about the parameter vector, and what it knows about its adversary. The minimizing player knows that the maximizing player has an advantage in terms of possessing the true value of the parameter. While the minimizing player is aware of this fact, since it is does not have the true value of the parameter vector, it can only solve the game under nominal conditions ($\rho = \hat{\rho}$). To construct safe trajectories (that minimize the chance of constraint violation) for the minimizing player under parametric uncertainties, we propose a modified

payoff function of the form

$$\tilde{\mathcal{J}}_{\hat{\rho}}(u,v) = \mathcal{J}_{\hat{\rho}}(u,v) + \int_{t_0}^{t_f} \|\text{vec } S^r(t)\|_Q^2 \, \mathrm{d}t, \tag{7.20}$$

where $\|w\|_Q = w^\top Q w$, and $Q \geq 0$. The subscript of $\tilde{\mathcal{J}}$ in Equation 7.20 denotes the parameter value that is used to evaluate the payoff function, given the control inputs u and v. The term $\int_{t_0}^{t_f} \|\text{vec } S^r(t)\|_Q^2$, hereafter referred to as the RCS regularizer, can be understood as a sensitivity-based measure that captures the risk of constraint violation. The minimizing player solves the non-zero-sum differential game

$$\min_{\substack{u\in\mathcal{U}\\ g^u\leq 0}} \tilde{\mathcal{J}}_{\hat{\rho}}(u,v) \max_{\substack{v\in\mathcal{V}\\ g^v\leq 0}} \mathcal{J}_{\hat{\rho}}(u,v). \tag{$\mathcal{G}^u$}$$

In ($\mathcal{G}^u$), the minimizing player, choosing the controller $u \in \mathcal{U}$, tries to minimize the augmented payoff function $\tilde{\mathcal{J}}_{\hat{\rho}}(u,v)$ that carries the sensitivity-based risk measure while satisfying its trajectory constraints ($g^u \leq 0$) under nominal conditions. Subsequently, the player choosing the controller $v \in \mathcal{V}$ tries to maximize the original payoff function while satisfying its constraints ($g^v \leq 0$) under nominal conditions. Accordingly, the minimizing player employs the state-feedback controller $\hat{u}^* : \mathbb{R}^n \to U$, where $(\hat{u}^*, \hat{v}^*)$ (with $\hat{v}^* : \mathbb{R}^n \to V$) is the solution to ($\mathcal{G}^u$).

As mentioned in section 7.1.A, the maximizing player is unaware of the modified payoff function of the minimizing player. However, it knows that the minimizing player only has information about the nominal value of the parameter $\hat{\rho}$. Consequently, the maximizing player solves the non-zero-sum differential game

$$\max_{\substack{v\in\mathcal{V}\\ g^v\leq 0}} \mathcal{J}_{\bar{\rho}}(u,v) \min_{\substack{u\in\mathcal{U}\\ g^u\leq 0}} \mathcal{J}_{\hat{\rho}}(u,v). \tag{$\mathcal{G}^v$}$$

Finally, the maximizing player employs the state-feedback controller $\bar{v}^* : \mathbb{R}^n \to V$, where $(\bar{u}^*, \bar{v}^*)$ (with $\bar{u}^* : \mathbb{R}^n \to U$) is the solution to $\mathcal{G}^v$.

7.3 Pursuit-Evasion with Asymmetric Information

In this section, we analyze a two-player pursuit-evasion problem with an uncertain dynamic obstacle using the approach presented in subsection 7.2.2. Here, the control u corresponds to the pursuer (minimizing player), and the control v corresponds to the evader (maximizing player).

Consider a one-pursuer one-evader problem with players' dynamics given by

$$\dot{x}^p = \mathrm{v}^p \cos u, \quad \dot{y}^p = \mathrm{v}^p \sin u, \tag{7.21}$$

$$\dot{x}^e = \mathrm{v}^e \cos v, \quad \dot{y}^e = \mathrm{v}^e \sin v, \tag{7.22}$$

where $(x^i, y^i) \in \mathbb{R}^2$, $i = \{p, e\}$, denote the positions of the pursuer (p) and the evader (e), v^p denotes the speed of the pursuer, and v^e denotes the speed of the evader. The initial positions of the agents are assumed to be given, and are denoted as $x^i(0) = a^i$, $y^i(0) = b^i$, $i = \{p, e\}$. In Equation 7.21 and Equation 7.22, u, v denote the headings (control inputs) of the pursuer and the evader, respectively. The pursuer's objective is to capture the evader by entering its capture zone, assumed here to be a disk of radius $\epsilon > 0$ centered at the instantaneous position of the evader, in finite time, whereas the evader's objective is to avoid capture indefinitely. At the same time, both agents have to avoid colliding with a dynamic circular obstacle of radius r_o, moving parallel to the y-axis with equations of motion

$$\dot{z} = \rho, \quad z(0) = d, \tag{7.23}$$

where (c, z) denotes the position of obstacle, and $|\rho|$ is its constant speed along the y-axis. For both agents, collision avoidance can be enforced by choosing the constraint functions

g^u, g^v of the form

$$g^u = r_o - \sqrt{(x^p - c)^2 + (y^p - z)^2} \leq 0, \tag{7.24}$$

$$g^v = r_o - \sqrt{(x^e - c)^2 + (y^e - z)^2} \leq 0, \tag{7.25}$$

$j = \{1, \ldots, N\}$. Finally, to ensure capturability of the evader, it is assumed that $\mathrm{v}^p > \mathrm{v}^e$. Also, it is assumed that each player can observe its instantaneous position and that of their adversary.

For this problem, we consider the speed of the obstacle to be the uncertain parameter i.e., the evader is aware of the actual speed of the obstacle ($\bar{\rho}$) while the pursuer has only the corresponding nominal value ($\hat{\rho}$) and cannot observe the obstacles' movements (until collision). To this end, safe trajectories for the pursuer can be generated using the regularizer introduced in subsection 7.2.2. Since the obstacles have simple dynamics, a closed form expression to the sensitivity of the pursuer's constraint function g^u with respect to the uncertain parameter ρ at time t is given by

$$S^u(t) = \frac{\partial g^u}{\partial \rho} = \frac{(z - y^p)t}{\sqrt{(x^p - c)^2 + (y^p - z)^2}}. \tag{7.26}$$

Owing to the intractability of solving the proposed pursuit-evasion problem analytically under the framework presented in subsection 7.2.2, the corresponding games ($\mathcal{G}^u$) and ($\mathcal{G}^v$) are solved numerically in a receding horizon fashion, presented in the following subsection.

7.3.1 Receding Horizon Control

In order to construct a receding horizon control (RHC) law for both players, we rewrite the dynamical equations in Equation 7.21-Equation 7.23 in discrete time as follows.

$$x^p_{k+1} = x^p_k + \mathrm{v}^p \cos u_k \Delta t, \quad y^p_{k+1} = y^p_k + \mathrm{v}^p \sin u_k \Delta t, \tag{7.27}$$

$$x^e_{k+1} = x^e_k + \mathrm{v}^e \cos v_k \Delta t, \quad y^e_{k+1} = y^e_k + \mathrm{v}^e \sin v_k \Delta t, \tag{7.28}$$

$$z_{k+1} = z_k + \rho \Delta t, \quad k = \{1, \ldots, N\}, \tag{7.29}$$

where the subscript k denotes the variable at the k^{th} time-step, and Δt is the length of time interval considered for discretization (assumed to be a constant). The initial conditions are given by $(x^i_0, y^i_0) = (a^i, b^i)$, $i = \{p, e\}$, $z_0 = d$. At each time-step, let $N\Delta t$ be the length of the horizon over which the games ($\mathcal{G}^u$) and ($\mathcal{G}^v$) are solved to obtain the open-loop controls of players.

Since the players follow an RHC law with fixed time horizon, we assume that the pursuer wants to minimize the relative distance at the end of the horizon, and vice versa for the evader. Consequently, at the k^{th} time-step in the RHC framework, the payoff function which the players want to optimize is given by

$$\mathcal{J}_\rho(\check{u}_{1:N}, \check{v}_{1:N}) = \sqrt{(x^p_{k+N} - x^e_{k+N})^2 + (y^p_{k+N} - y^e_{k+N})^2}, \tag{7.30}$$

where $\check{u}_{1:N},\ \check{v}_{1:N}$ are the control inputs of the players for the optimization as per the RHC law. Note that the time interval Δt is assumed to be constant and hence, the augmented payoff function along with the RCS regularizer the pursuer minimizes at the k^{th} time-step is given by

$$\tilde{\mathcal{J}}_{\hat{\rho}}(\check{u}_{1:N}, \check{v}_{1:N}) = \mathcal{J}_{\hat{\rho}} + \sum_{i=1}^{N} \|\mathrm{vec}\ S^r_{k+i}\|^2_Q, \tag{7.31}$$

where S^r_{k+i} is the RCS at the i^{th} step in the time horizon. The RHC algorithm the agents use to obtain their respective control inputs at the k^{th} time-step is given in Algorithm 4. Assuming that there exists solutions for both games ($\mathcal{G}^u$) and ($\mathcal{G}^v$), the well-known Gauss-Seidel method is employed to obtain the fixed points for both games [180]. The constraints g^u and g^v are imposed at discrete time-steps for collision avoidance in the optimization problems stated in Algorithm 4.

Algorithm 4 Receding horizon control law for the players

procedure RHC(N, Δt, x^p_k, x^e_k)
  Initialize with $[\check{v}_{1:N}]_0$, $i \leftarrow 1$ ▷ Pursuer's game ($\mathcal{G}^u$)
  repeat
    $[\check{u}_{1:N}]_i \leftarrow \underset{\check{u}_{1:N},\ g^u \leq 0}{\arg\min}\ \tilde{J}_{\hat{\rho}}(\check{u}_{1:N}, [\check{v}_{1:N}]_{i-1})$
    $[\check{v}_{1:N}]_i \leftarrow \underset{\check{v}_{1:N},\ g^v \leq 0}{\arg\min}\ J_{\hat{\rho}}([\check{u}_{1:N}]_i, \check{v}_{1:N})$
    $i \leftarrow i + 1$
  until convergence
  $u_k \leftarrow [\check{u}_1]_{i-1}$
  Initialize with $[\check{v}_{1:N}]_0$, $i \leftarrow 1$ ▷ Evader's game ($\mathcal{G}^v$)
  repeat
    $[\check{u}_{1:N}]_i \leftarrow \underset{\check{u}_{1:N},\ g^u \leq 0}{\arg\min}\ J_{\hat{\rho}}(\check{u}_{1:N}, [\check{v}_{1:N}]_{i-1})$
    $[\check{v}_{1:N}]_i \leftarrow \underset{\check{v}_{1:N},\ g^v \leq 0}{\arg\min}\ J_{\bar{\rho}}([\check{u}_{1:N}]_i, \check{v}_{1:N})$
    $i \leftarrow i + 1$
  until convergence
  $v_k \leftarrow [\check{v}_1]_{i-1}$
  return $u_k,\ v_k$
end procedure

7.3.2 Simulations

In this subsection, we present the simulation results for the pursuit-evasion game with one uncertain dynamic obstacle. The simulations are performed in MATLAB using its in-built function `fmincon` in conjunction with YALMIP [181] for the optimization problems shown in Algorithm 4. The simulation parameters corresponding to the players and the obstacle are chosen to be $\epsilon = 0.3$, $\mathrm{v}^p = 1$, $\mathrm{v}^e = 0.6$, $\hat{\rho} = 0.25$, $(a^p, b^p) = (0, 0)$, $(a^e, b^e) = (3, 0)$, $(c, d) = (2, 1.15)$, $r_o = 0.5$. The variation in the obstacle's speed is chosen to be

$\bar{\rho} - \hat{\rho} = 0.1$. The parameters corresponding to the RHC control law are set to $N = 10$, $\Delta t = 0.1$. The convergence criterion to obtain the equilibrium points for the games $\mathcal{G}^u$ and $\mathcal{G}^v$, as per the Algorithm 4, is chosen to be $\|[\check{u}_{1:N}]_{i-1} - [\check{u}_{1:N}]_{i-1}\|, \|[\check{v}_{1:N}]_{i-1} - [\check{v}_{1:N}]_{i-1}\| \leq 10^{-4}$

First, we examine the case where the pursuer does not desensitize, which can be captured by setting the penalty on the relevant constraint sensitivity (RCS) to be $Q = 0$. In this case, the pursuer ends up colliding with the obstacle, and the result is presented in Figure 7.1. The filled grey circle denotes the obstacle's instantaneous position (not observed by the pursuer) while the dotted circle represents its initial position. The blue line denotes the trajectory traced by the pursuer with the markers denoting the initial and the instantaneous position of the pursuer. Similarly, red corresponds to the evader. The green curves denote the instantaneous N-step look ahead trajectories for both players based on the RHC control law. The collision in this case can be attributed to the fact that the pursuer considers just the nominal model of the obstacle, which cannot be observed until collision, and it does not consider "the risk" associated to it.

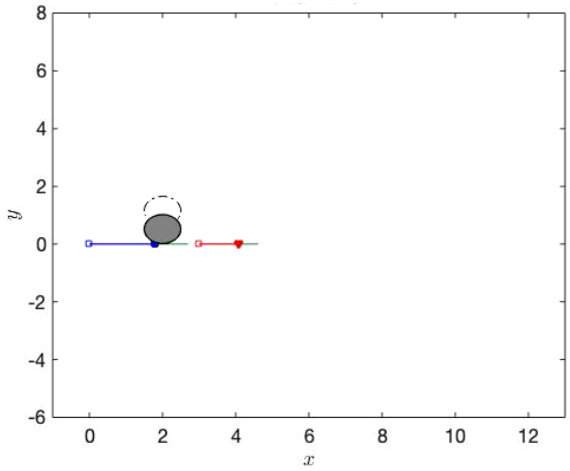
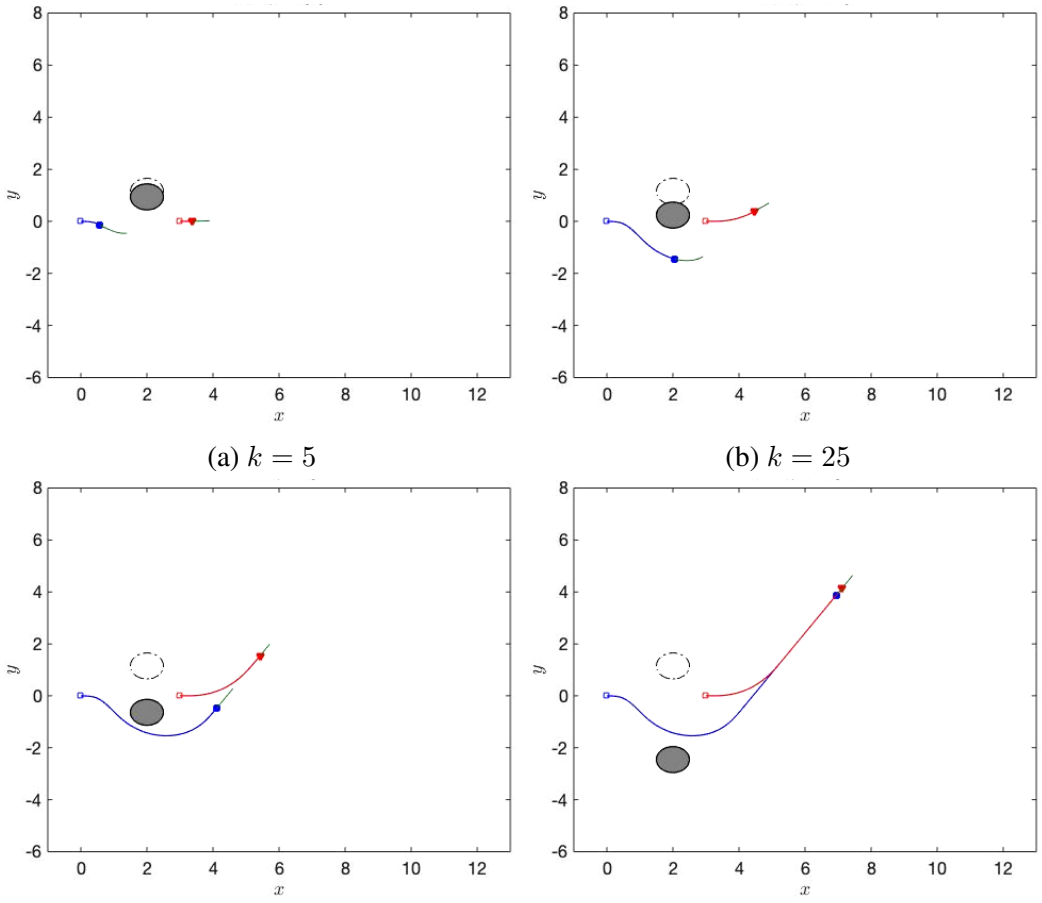


Figure 7.1: An instance where the pursuer does not penalize the RCS ($Q = 0$) that leads to collision

Figure 7.2 presents the results obtained for the case where $Q = 1$. In this case, the pursuer solves a multi-objective optimization that minimizes the risk of collision, captured



(a) $k = 5$ (b) $k = 25$

(c) $k = 50$ (d) $k = 102$

Figure 7.2: Simulation results for the instance where the pursuer penalizes the RCS cost with $Q = 1$, and successfully captures the evader while avoiding the obstacle

using the norm of RCS, along with the relative distance at the final time-step. This is evident from the fact that the pursuer's instantaneous look-ahead trajectory (the green curve) takes the pursuer away from the obstacle, as opposed to pursuing the evader along its line-of-sight, even when the pursuer and the look-ahead trajectory is sufficiently away from the obstacle. As a result, the pursuer maneuvers around the obstacle while maintaining a safe distance from it. The pursuer is also successful in capturing the evader.

It has to be noted that the pursuer's trajectory in Figure 7.2 is efficient in the sense that it avoids collisions for a large range of variations of the obstacle speed. As observed in chapter 6, higher penalty on the RCS norm will provide a trajectory that is prone to lesser collisions under parametric variations while trading off capture time.

CHAPTER 8
CONCLUSIONS

8.1 Summary

In this dissertation, we presented solutions to pursuit-evasion problems containing multiple agents and uncertainties. First, we considered problem formulations that involve multiple pursuers and multiple evaders with deterministic dynamics. Since multi-pursuer multi-evader problems are known to be intractable, heuristics based on the geometry of the game are employed to obtain task-allocation algorithms that are computationally efficient. This is achieved by first analyzing pursuit-evasion problems involving two pursuers and one evader. In chapter 2, two pursuers/one evader problems are analyzed in two different scenarios, assuming both pursuers to be superior to the evader in terms of their speed capabilities. The first scenario involves pursuers that have access to information of the evader's position and velocity and use this information to follow a constant bearing strategy. In the second scenario, the pursuers have access to the evader's position only and hence, they follow a pure pursuit strategy. The time optimal evading strategy is identified in both scenarios. Since obtaining a closed-form solution in the second scenario is elusive at this point, a competitive suboptimal strategy that can be practically implemented is identified and is compared against the optimal strategy. The regions of non-degeneracy are used to investigate the utility of employing two pursuers to more efficiently capture the evader. If the initial positions of the players are such that the problem is degenerate, then one of the pursuers does not play any role in the game, and the optimal evading strategy is pure evasion from the other pursuer. Optimal evading strategies against relay pursuit are also investigated by keeping one pursuer stationary. The results presented in this chapter provides a potential framework for solving larger classes of multi-player time-optimal pursuit-evasion

games under different information structures.

In chapter 3, under the assumption that the pursuers are faster than the evader(s), and that they follow either a constant bearing (CB) or a pure pursuit (PP) strategy, workable solutions for multi-pursuer single-evader (MPSE) and multi-pursuer multi-evader (MPME) problems are provided. In both CB and PP cases, it has been established that the optimal evading strategy in the MPSE setting depends only on those pursuers that capture the evader simultaneously. Using this insight, a dynamic allocation algorithm for the pursuers, which is independent of the evader's strategy, has been proposed to solve the MPSE problem. The proposed algorithm is based on the notion of active/redundant pursuers and their classification using the Apollonius cycles (for the case of CB) or the Apollonius curves (for the case of PP). The algorithm is further extended to solve MPME problems for any number of pursuers and evaders. These algorithms ensure capture of all the evaders either in an MPSE or an MPME setting in finite time. Subsequently, the proposed solution techniques are extended to problems with heterogeneous teams of pursuers and evaders.

Second, we considered uncertainties that are exogenous and stochastic in nature. In chapter 4, using the theory of general constrained games, the problem of steering a Gaussian in adversarial scenarios is studied. The problem is posed from a perspective of the player that desires to drive the distribution to a given terminal Gaussian while minimizing a quadratic cost. The player that tries to maximize the cost is assumed to be indifferent to the terminal constraint. It is shown that the game need not have a saddle point equilibrium. Subsequently, we obtained necessary conditions for the controller to drive the mean to the specified value in the upper game. The covariance steering problem is solved numerically using the well-known Jacobi procedure. The proposed approach is applied on the missile endgame guidance problem.

In chapter 5, a novel approach to addressing pursuit-evasion problems under external stochastic flow fields is presented. The players' nominal trajectories are obtained using forward reachability analysis while ignoring the diffusion part of the flow field. The nom-

inal solution thus obtained is time-optimal for the players under deterministic conditions. Assuming a linear feedback control strategy, a chance-constrained covariance game is constructed around the nominal solution. The proposed covariance steering game involves optimizing over the value of the smallest relative distance that can be achieved with high probability. The pursuer tries to minimize this value while the evader tries to maximize it by equivalently optimizing over the covariance of the relative distance. The proposed approach is tested on realistic linear and nonlinear flow fields. Numerical simulations suggest that the pursuer can effectively steer the game towards capture while controlling the covariance of the relative distance.

Finally, we considered uncertainties that are parametric in nature. In chapter 6, the associated optimal control formulation is analyzed initially using sensitivity functions. Various schemes for trajectory/state desensitization are proposed which compete with the existing approaches in terms of the computational complexity, and it is realized that the sensitivity function is more tractable compared to the Seywald's sensitivity matrix. The proposed schemes are demonstrated using fixed final time, and minimum time Zermelo's optimal control problems. Subsequently, a sensitivity function-based regularizer is introduced to obtain conservative solutions that avoid constraint violation under parametric uncertainties in optimal control problems. Using the fact that collision avoidance can be expressed as a state constraint, the approach is applied for path planning problems involving dynamic uncertain obstacles. The proposed regularizer is first analyzed on simple problems to study its characteristics and to identify its limitations. It is observed that the form of the constraint function used to construct the regularizer affects the behavior of the trajectories. The results on environments with as many as ten dynamic obstacles indicate that safety can be enhanced with an acceptable trade-off in optimality.

In chapter 7, the sensitivity function based approach is extended to game-theoretic formulations to address planning in adversarial environments with asymmetric information. The uncertainty associated to the asymmetric information is assumed to be parametric in

nature. To this end, penalizing the norm of the relevant constraint sensitivity, which can be understood as a form of risk measure, is shown to provide safe trajectories that reduce the risk of constraint violations. In a two-player pursuit-evasion game in which the pursuer has information only about the nominal motion model of an uncertain obstacle, it is demonstrated that the pursuer can generate trajectories that can capture the evader while successfully avoiding the obstacle under variations in its nominal motion model.

8.2 Recommendations for Future Research

The work presented in chapter 2 provides interesting directions for future research. Algorithm 1 and the Apollonius allocation (A2) algorithm are implemented at every time-step to identify active pursuers for each evader. The computational requirements associated to the implementation of these algorithms can be improved by having an estimate of when the assignment can change to avoid unnecessary calculations at every time-step. The A2 algorithm identifies the redundant pursuers to save resources in terms of fuel. However, it does not consider fuel limitations posed for a single vehicle, as is the case in real-world scenarios. Therefore, it is been observed that a single pursuer can be assigned to the active duty of pursuing evaders for a considerably longer time compared to the rest of the pursuers in its team. To this end, the proposed algorithms for multi-pursuer multi-evader (MPME) problem can be extended to include constraints on the total time for which a pursuer can remain active. A potential solution approach to this problem formulation can involve employing the pursuers that are found to be redundant as per the A2 algorithm.

Another interesting research direction is to consider non-holonomic dynamics that imposes turn-radius constraints for players in MPME problems [189, 190]. One of the promising approaches to solve this problem can involve extending the notion of Apollonius curves to account for the players' turn-radius constraints. The concept of an evader's safe-reachable area in pursuit-evasion problems with non-holonomic systems can be utilized in this regard [191]. The boundary of an evader's safe reachable set can be considered as the general-

ization of the Apollonius circle. It has to be noted that the Apollonius circle is expressed in closed-form as a function of the players' instantaneous positions and speed capabilities. However, the safe reachable set of an evader, which is assumed to be a Dubins vehicle, in a pursuit-evasion game can only be obtained numerically. In MPME scenarios, obtaining such safe reachable sets for every pursuer-evader pair can be computationally challenging. Once the boundary of the safe reachable set associated to a single pursuer-evader pair can be obtained efficiently, the next challenge is to define a closed curve around each evader that is analogous to the Apollonius boundary to identify its active pursuers. In order to address airspace security using the solution approaches proposed in this thesis, it is important to extend them to scenarios with closed domains. Since in the case of guarding a territory, it is common to assume that successful evasion involves leaving the game space by being able to cross the boundary of the closed domain.

In chapter 4 and chapter 5, it is assumed that each player observes the instantaneous positions of the players and the control input of their adversary at the previous time-step. This is a very strong assumption, and future work can consider output-feedback structure, that is closer to real-world situations. An observer for each player can be constructed to estimate the state of their adversary while employing an output-feedback control. As demonstrated in [72], the players can choose to increase the uncertainty in their state along with optimizing the relative distance. This can prove to be an efficient strategy for the evader, as shown in chapter 5, when it is slower compared to the pursuer.

In the covariance steering game from chapter 4, it is assumed that the maximizing player is indifferent to the minimizing player's terminal constraint. The terminal constraint is a form of capture condition for pursuit-evasion problems. Since the evader would ideally want to avoid capture indefinitely, the case in which the maximizing player (evader) is adversarial to the terminal constraint is worth analyzing. In the corresponding mean steering game, a condition similar to the relative controllability matrix being full rank can be obtained for both players. Such a result will establish the fact regarding who can efficiently

drive the state to their desired terminal condition.

Extension to multi-player stochastic pursuit-evasion problems is an important area of research. The forward reachability based approach for deterministic pursuit-evasion, discussed in chapter 5, is developed for multi-pursuer single-evader (MPSE) games [66]. Therefore, in MPSE scenarios, the nominal trajectories of the players can be obtained from the forward reachability analysis, as discussed in Ref. [66]. First, a chance-constrained covariance game involving multiple pursuers and single evader can be formulated to obtain feedback strategies for all players. The interesting aspect of this research direction is to analyze the aspect of cooperation among the pursuers. The payoff function for each pursuer can be designed such that the pursuers capture the evader with high probability in a decentralized and non-cooperative fashion. Alternatively, there can be a payoff function that is common to all pursuers such that the evader can be captured with high probability in a cooperative fashion. The importance of cooperation can be quantified by comparing the performance under the aforementioned payoff functions. Also, it is important to analyze the existence and uniqueness of the equilibrium solution in these covariance steering games

One of the challenging research directions is to develop computationally efficient algorithms to solve these games numerically. The Gauss-Seidel and Jacobi iterative techniques require the payoff function to be convex to guarantee the existence of a stable equilibrium, and subsequently, convergence of these algorithms. In most cases, the pursuit-evasion problem formulations are non-convex, and convexification methods can be explored to address this problem [192, 193, 194].

Finally, desensitized optimal control (DOC) and its game-theoretic extension provide promising research directions. The idea of DOC can be analyzed from the perspective of risk-sensitive optimal control. Assuming that the parameter is a Gaussian random variable $\rho \sim \mathcal{N}(\hat{\rho}, P)$, the expected output variation can be approximated using sensitivity functions

as

$$\mathbb{E}[\delta y \delta y^\top] \approx \mathrm{tr}[Y S P S^\top Y^\top], \tag{8.1}$$

where y is the output vector, Y is the Jacobian of the output function with respect to the state vector x, and S is the sensitivity function. It has been shown that by penalizing the expected output variations, as opposed to the norm of the sensitivity matrix, open-loop trajectories that are robust to parametric variations can be generated in the case of hypersonic vehicle trajectory optimization [195]. This approach is very much inline with the concept of risk-sensitive optimal control, where the variance of the cost function is penalized along with the expected cost to ensure robustness.

In chapter 6, the sensitivity function based regularization techniques ensure robustness of the open-loop trajectories under parametric variations. The approach can be extended to include feedback control design into the DOC framework by considering the sensitivity equation of the closed-loop trajectory, given by

$$\dot{S} = (A + CK)S + B, \quad S(t_0) = 0, \tag{8.2}$$

where K is the closed-loop gain that is obtained under the assumption of linear feedback control $\Delta u = K \Delta x$. Here, A, B, C denote the Jacobian of the state dynamics with respect to state, parameter, and control, respectively (evaluating along the nominal trajectory). Previously work assumed that the gain K to be a constant [113], and was independent of the nominal control that is obtained while regularizing the sensitivity function. Preliminary research suggests that the gain K can be driven by the choice of the nominal control by augmenting the dynamics of the Riccati equation to the set of constraints. This will most likely provide a coupled nominal and feedback control with an acceptable trade-off between the corresponding performance metrics.

Finally, the problem formulation that is discussed in chapter 7 can be extended to in-

clude the aspect of deception by the player who has an information advantage. For example in the pursuit-evasion game, discussed in chapter 7, it is worth analyzing whether the relevant constraint sensitivity (RCS) regularizer is efficient against active deception by the evader. Assuming the evader knows that the pursuer has only the nominal model of the obstacle, and both players follow a position-feedback-based strategy, the evader can maximize the chance of pursuer colliding the obstacle by modifying its own payoff function. This will involve luring the pursuer closer to the obstacle due to its lack of information regarding the parametric variations, and based on the pursuer's feedback strategy.

REFERENCES

[1] Y. Huang, S. J. Thomson, W. C. Hoffmann, Y. Lan, and B. K. Fritz, "Development and prospect of unmanned aerial vehicle technologies for agricultural production management," *International Journal of Agricultural and Biological Engineering*, vol. 6, no. 3, pp. 1–10, 2013.

[2] F. Nex and F. Remondino, "UAV for 3D mapping applications: A review," *Applied Geomatics*, vol. 6, no. 1, pp. 1–15, 2014.

[3] C. Yuan, Y. Zhang, and Z. Liu, "A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques," *Canadian Journal of Forest Research*, vol. 45, no. 7, pp. 783–792, 2015.

[4] C. A. F. Ezequiel, M. Cua, N. C. Libatique, G. L. Tangonan, R. Alampay, R. T. Labuguen, C. M. Favila, J. L. E. Honrado, V. Canos, C. Devaney, *et al.*, "UAV aerial imaging applications for post-disaster assessment, environmental management and infrastructure development," in *International Conference on Unmanned Aircraft Systems*, IEEE, Orlando, FL, 2014, pp. 274–283.

[5] B. C. Beckman, M. Haskin, M. Rolnik, and Y. Vule, *Maneuvering a package following in-flight release from an unmanned aerial vehicle (UAV)*, Patent No.: US 9,567,081 B1, Feb. 2017.

[6] T. Mylvaganam, M. Sassano, and A. Astolfi, "A constructive differential game approach to collision avoidance in multi-agent systems," in *American Control Conference*, IEEE, 2014, pp. 311–316.

[7] V. Isler, D. Sun, and S. Sastry, "Roadmap based pursuit-evasion and collision avoidance," in *Robotics: Science and Systems*, vol. 1, Cambridge, United States, 2005, pp. 257–264.

[8] I. Exarchos, P. Tsiotras, and M. Pachter, "UAV collision avoidance based on the solution of the suicidal pedestrian differential game," in *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2016-2100, San Diego, CA, 2016.

[9] J. Las Fargeas, P. Kabamba, and A. Girard, "Cooperative surveillance and pursuit using unmanned aerial vehicles and unattended ground sensors," *Sensors*, vol. 15, no. 1, pp. 1365–1388, 2015.

[10] R. Isaacs, *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*. Mineola, NY: Dover Publications, Inc., 1999, Chapter 6.

[11] S. S. Kumkov, S. Le Ménec, and V. S. Patsko, "Zero-sum pursuit-evasion differential games with many objects: Survey of publications," *Dynamic Games and Applications*, vol. 7, no. 4, pp. 609–633, 2017.

[12] H. J. Kelley, "Some aspects of two-on-one pursuit/evasion," *Automatica*, vol. 9, no. 3, pp. 403–404, 1973.

[13] S. Bhattacharya and S. Hutchinson, "Approximation schemes for two-player pursuit evasion games with visibility constraints," in *Robotics: Science and Systems*, vol. 1, Zurich, Switzerland, 2008, pp. 81–88.

[14] M. Foley and W. Schmitendorf, "A class of differential games with two pursuers versus one evader," *IEEE Transactions on Automatic Control*, vol. 19, no. 3, pp. 239–243, 1974.

[15] S. Le Ménec, "Linear differential game with two pursuers and one evader," in *Advances in Dynamic Games: Theory, Applications, and Numerical Methods for Differential and Stochastic Games*, M. Breton and K. Szajowski, Eds. Boston, MA: Birkhäuser, 2011, pp. 209–226.

[16] S. A. Ganebny, S. S. Kumkov, S. Le Ménec, and V. S. Patsko, "Study of linear game with two pursuers and one evader: Different strength of pursuers," in *Advances in Dynamic Games: Theory, Applications, and Numerical Methods for Differential and Stochastic Games*, P. Cardaliaguet and R. Cressman, Eds. Boston, MA: Birkhäuser, 2013, pp. 269–292.

[17] S. S. Kumkov, S. Le Ménec, and V. S. Patsko, "Level sets of the value function in differential games with two pursuers and one evader. interval analysis interpretation," *Mathematics in Computer Science*, vol. 8, no. 3, pp. 443–454, 2014.

[18] A. Y. Levchenkov and A. G. Pashkov, "Differential game of optimal approach of two inertial pursuers to a noninertial evader," *Journal of Optimization Theory and Applications*, vol. 65, no. 3, pp. 501–518, 1990.

[19] P. Hagedorn and J. V. Breakwell, "A differential game with two pursuers and one evader," *Journal of Optimization Theory and Applications*, vol. 18, no. 1, pp. 15–29, 1976.

[20] W. Sun and P. Tsiotras, "An optimal evader strategy in a two-pursuer one-evader problem," in *IEEE Conference on Decision and Control*, Los Angeles, CA, Dec. 2014, pp. 4266–4271.

[21] B. Pshenichnyi, "Simple pursuit by several objects," *Cybernetics and Systems Analysis*, vol. 12, no. 3, pp. 484–485, 1976.

[22] A. Blagodatskikh, "Simultaneous multiple capture in a simple pursuit problem," *Journal of Applied Mathematics and Mechanics*, vol. 73, no. 1, pp. 36–40, 2009.

[23] F. Chernous'ko, "A problem of evasion from many pursuers," *Journal of Applied Mathematics and Mechanics*, vol. 40, no. 1, pp. 11–20, 1976.

[24] G. I. Ibragimov, "Optimal pursuit with countably many pursuers and one evader," *Differential Equations*, vol. 41, no. 5, pp. 627–635, 2005.

[25] G. I. Ibragimov, M. Salimi, and M. Amini, "Evasion from many pursuers in simple motion differential game with integral constraints," *European Journal of Operational Research*, vol. 218, no. 2, pp. 505–511, 2012.

[26] J. S. Jang and C. J. Tomlin, "Control strategies in multi-player pursuit and evasion game," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, AIAA Paper 2005-6239, San Francisco, CA, 2005.

[27] V. Zak, "On a problem of evading many pursuers," *Journal of Applied Mathematics and Mechanics*, vol. 43, no. 3, pp. 492–501, 1979.

[28] D. W. Oyler, P. T. Kabamba, and A. R. Girard, "Pursuit–evasion games in the presence of obstacles," *Automatica*, vol. 65, pp. 1–11, 2016.

[29] M. V. Ramana and M. Kothari, "A cooperative pursuit-evasion game of a high speed evader," in *IEEE Conference on Decision and Control*, Osaka, Japan, 2015, pp. 2969–2974.

[30] ——, "A cooperative pursuit strategy for a high speed evader," in *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2016-2103, San Diego, CA, 2016.

[31] M. V. Ramana and M. Kothari, "Pursuit strategy to capture high-speed evaders using multiple pursuers," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 1, pp. 139–149, 2017.

[32] ——, "Pursuit-evasion games of high speed evader," *Journal of Intelligent & Robotic Systems*, vol. 85, no. 2, pp. 293–306, 2017.

[33] T. R. Vechalapu, "A trapping pursuit strategy for capturing a high speed evader," in *AIAA Scitech Forum*, AIAA Paper 2020-2069, Orlando, FL, 2020.

[34] S. D. Bopardikar, F. Bullo, and J. P. Hespanha, "A cooperative homicidal chauffeur game," *Automatica*, vol. 45, no. 7, pp. 1771–1777, 2009.

[35] M. Pittsyk and A. Chikrii, "On a group pursuit problem," *Journal of Applied Mathematics and Mechanics*, vol. 46, no. 5, pp. 584–589, 1982.

[36] W. Chodun, "Differential games of evasion with many pursuers," *Journal of Mathematical Analysis and Applications*, vol. 142, no. 2, pp. 370–389, 1989.

[37] A. I. Blagodatskikh, "Group pursuit in Pontryagin's nonstationary example," *Differential Equations*, vol. 44, no. 1, pp. 40–46, 2008.

[38] J. P. Hespanha, H. J. Kim, and S. Sastry, "Multiple-agent probabilistic pursuit-evasion games," in *IEEE Conference on Decision and Control*, vol. 3, Phoenix, AZ, 1999, pp. 2432–2437.

[39] E. Bakolas and P. Tsiotras, "On the relay pursuit of a maneuvering target by a group of pursuers," in *IEEE Conference on Decision and Control and European Control Conference*, Orlando, FL, Dec. 2011, pp. 4270–4275.

[40] E. Bakolas and P. Tsiotras, "Relay pursuit of a maneuvering target using dynamic Voronoi diagrams," *Automatica*, vol. 48, no. 9, pp. 2213–2220, 2012.

[41] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. 1947.

[42] J. Ge, L. Tang, J. Reimann, and G. Vachtsevanos, "Aiaa guidance, navigation, and control conference and exhibit," in *Guidance, Navigation, and Control Conference and Exhibit*, AIAA Paper 2006-6786, Keystone, CO, 2006.

[43] D. Li, J. B. Cruz, G. Chen, C. Kwan, and M.-H. Chang, "A hierarchical approach to multi-player pursuit-evasion differential games," in *IEEE Conference on Decision and Control and European Control Conference*, Seville, Spain, 2005, pp. 5674–5679.

[44] S. Jin and Z. Qu, "A heuristic task scheduling for multi-pursuer multi-evader games," in *International Conference on Information and Automation*, IEEE, Shenzhen, China, 2011, pp. 528–533.

[45] A. Antoniades, H. J. Kim, and S. Sastry, "Pursuit-evasion strategies for teams of multiple agents with incomplete information," in *IEEE Conference on Decision and Control*, Maui, HI, 2003, pp. 756–761.

[46] D. M. Stipanović, A. Melikyan, and N. Hovakimyan, "Guaranteed strategies for nonlinear multi-player pursuit-evasion games," *International Game Theory Review*, vol. 12, no. 1, pp. 1–17, 2010.

[47] ——, "Some sufficient conditions for multi-player pursuit-evasion games with continuous and discrete observations," in *Advances in Dynamic Games and Their Applications: Analytical and Numerical Developments*, O. Pourtallier, V. Gaitsgory, and P. Bernhard, Eds. Boston: Birkhäuser Boston, 2009, pp. 1–13.

[48] A. Pierson, Z. Wang, and M. Schwager, "Intercepting rogue robots: An algorithm for capturing multiple evaders with multiple pursuers," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 530–537, 2017.

[49] D. Shishika and V. Kumar, "Local-game decomposition for multiplayer perimeter-defense problem," in *IEEE Conference on Decision and Control*, Miami, FL, 2018, pp. 2093–2100.

[50] D. Shishika, J. Paulos, and V. Kumar, "Cooperative team strategies for multi-player perimeter-defense games," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2738–2745, 2020.

[51] D. Shishika and V. Kumar, "A review of multi agent perimeter defense games," in *Decision and Game Theory for Security*, Q. Zhu, J. S. Baras, R. Poovendran, and J. Chen, Eds., Cham: Springer International Publishing, 2020, pp. 472–485.

[52] J. K. Gupta, M. Egorov, and M. Kochenderfer, "Cooperative multi-agent control using deep reinforcement learning," in *Autonomous Agents and Multiagent Systems*, Springer, 2017, pp. 66–83.

[53] M. Hüttenrauch, S. Adrian, G. Neumann, *et al.*, "Deep reinforcement learning for swarm systems," *Journal of Machine Learning Research*, vol. 20, no. 54, pp. 1–31, 2019.

[54] A. T. Bilgin and E. Kadioglu-Urtis, "An approach to multi-agent pursuit evasion games using reinforcement learning," in *International Conference on Advanced Robotics*, IEEE, Istanbul, Turkey, 2015, pp. 164–169.

[55] S. F. Desouky and H. M. Schwartz, "Learning in n-pursuer n-evader differential games," in *IEEE International Conference on Systems, Man and Cybernetics*, Istanbul, Turkey, 2010, pp. 4069–4074.

[56] Y. Wang, L. Dong, and C. Sun, "Cooperative control for multi-player pursuit-evasion games with reinforcement learning," *Neurocomputing*, vol. 412, pp. 101–114, 2020.

[57] S. Li, Y. Wu, X. Cui, H. Dong, F. Fang, and S. Russell, "Robust multi-agent reinforcement learning via minimax deep deterministic policy gradient," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, Honolulu, Hawaii, 2019, pp. 4213–4220.

[58] R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," in *31st Conference on Neural Information Processing Systems (NIPS)*, Long Beach, CA, 2017, pp. 6382–6393.

[59] C. de Souza, R. Newbury, A. Cosgun, P. Castillo, B. Vidolov, and D. Kulić, "Decentralized multi-agent pursuit using deep reinforcement learning," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4552–4559, 2021.

[60] L. Busoniu, R. Babuska, and B. De Schutter, "A comprehensive survey of multiagent reinforcement learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 2, pp. 156–172, 2008.

[61] Y. Yavin, "Stochastic pursuit-evasion differential games in the plane," *Journal of optimization theory and applications*, vol. 50, no. 3, pp. 495–523, 1986.

[62] R. Vidal, O. Shakernia, H. J. Kim, D. H. Shim, and S. Sastry, "Probabilistic pursuit-evasion games: Theory, implementation, and experimental evaluation," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 5, pp. 662–669, 2002.

[63] P. Kachroo, S. A. Shedied, and H. Vanlandingham, "Pursuit evasion: The herding noncooperative dynamic game-The stochastic model," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 32, no. 1, pp. 37–42, 2002.

[64] D. Li, J. B. Cruz Jr, and C. J. Schumacher, "Stochastic multi-player pursuit–evasion differential games," *International Journal of Robust and Nonlinear Control*, vol. 18, no. 2, pp. 218–247, 2008.

[65] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin, "A time-dependent hamilton-jacobi formulation of reachable sets for continuous dynamic games," *IEEE Transactions on Automatic Control*, vol. 50, no. 7, pp. 947–957, 2005.

[66] W. Sun, P. Tsiotras, T. Lolla, D. N. Subramani, and P. F. Lermusiaux, "Multiple-pursuer/one-evader pursuit–evasion game in dynamic flowfields," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 7, pp. 1627–1637, 2017.

[67] H. Nabi, T. Lombaerts, Y. Zhang, E. van Kampen, Q. Chu, and C. C. de Visser, "Effects of structural failure on the safe flight envelope of aircraft," *Journal of Guidance, Control, and Dynamics*, vol. 41, no. 6, pp. 1257–1275, 2018.

[68] A. P. Vinod, B. HomChaudhuri, and M. M. Oishi, "Forward stochastic reachability analysis for uncontrolled linear systems using fourier transforms," in *International Conference on Hybrid Systems: Computation and Control*, ACM, Pittsburgh, PA, 2017, pp. 35–44.

[69] A. P. Vinod, B. HomChaudhuri, C. Hintz, A. Parikh, S. P. Buerger, M. M. Oishi, G. Brunson, S. Ahmad, and R. Fierro, "Multiple pursuer-based intercept via forward stochastic reachability," in *American Control Conference*, IEEE, Milwaukee, WI, 2018, pp. 1559–1566.

[70] J. Bertram and P. Wei, "Distributed computational guidance for high-density urban air mobility with cooperative and non-cooperative collision avoidance," in *AIAA Scitech Forum*, AIAA Paper 2020-1371, Orlando, FL, 2020.

[71] A. Jain, D. Gueho, and P. Singla, "A computationally efficient approach for stochastic reachability set analysis," in *AIAA Scitech Forum*, AIAA Paper 2020-0851, Orlando, FL, 2020.

[72] J. L. Speyer, "A stochastic differential game with controllable statistical parameters," *IEEE Transactions on Systems Science and Cybernetics*, vol. 3, no. 1, pp. 17–20, 1967.

[73] C. T. Leondes and B. Mons, "On-line solution of a stochastic pursuit-evasion game," *Journal of Optimization Theory and Applications*, vol. 28, no. 3, pp. 411–428, 1979.

[74] D. Castañón and M. Athans, "On stochastic dynamic Stackelberg strategies," *Automatica*, vol. 12, no. 2, pp. 177–183, 1976.

[75] P. Kumar and J. V. Schuppen, "On Nash equilibrium solutions in stochastic dynamic games," *IEEE Transactions on Automatic Control*, vol. 25, no. 6, pp. 1146–1149, 1980.

[76] K. B. Bley and E. B. Stear, "Discrete stochastic differential games," in, ser. Advances in Control Systems, C. Leondes, Ed., vol. 8, Elsevier, 1971, pp. 89–140.

[77] Y. Yavin, "The numerical solution of three stochastic differential games," *Computers & Mathematics with Applications*, vol. 10, no. 3, pp. 207–234, 1984.

[78] ——, "A stochastic two-target pursuit-evasion differential game with three players moving in a plane," in *Pursuit-Evasion Differential Games*, ser. International Series in Modern Applied Mathematics and Computer Science, Amsterdam: Pergamon, 1987, pp. 141–149.

[79] P. Bernhard and A. .-L. Colomb, "Saddle point conditions for a class of stochastic dynamical games with imperfect information," *IEEE Transactions on Automatic Control*, vol. 33, no. 1, pp. 98–101, 1988.

[80] A. Gupta, A. Nayyar, C. Langbort, and T. Başar, "Common information based Markov perfect equilibria for linear-Gaussian games with asymmetric informa-

tion," *SIAM Journal on Control and Optimization*, vol. 52, no. 5, pp. 3228–3260, 2014.

[81] T. Başar, "Stochastic differential games and intricacy of information structures," in *Dynamic Games in Economics*, J. Haunschmied, V. M. Veliov, and S. Wrzaczek, Eds. Berlin, Heidelberg: Springer, 2014, pp. 23–49.

[82] A. Swarup and J. L. Speyer, "Linear-Quadratic-Gaussian differential games with different information patterns," in *IEEE Conference on Decision and Control*, vol. 4, Maui, HI, 2003, pp. 4146–4151.

[83] J. W. Clemens and J. L. Speyer, "The LQG game with nonclassical information pattern using a direct solution method," in *American Control Conference*, IEEE, Seattle, WA, 2017, pp. 418–423.

[84] I. Rhodes and D. Luenberger, "Differential games with imperfect state information," *IEEE Transactions on Automatic Control*, vol. 14, no. 1, pp. 29–38, 1969.

[85] C.-Y. Chong and M. Athans, "On the stochastic control of linear systems with different information sets," *IEEE Transactions on Automatic Control*, vol. 16, no. 5, pp. 423–430, 1971.

[86] T. Rajpurohit, W. M. Haddad, and W. Sun, "Stochastic differential games and inverse optimal control and stopper policies," *International Journal of Control*, vol. 92, no. 7, pp. 1692–1706, 2019.

[87] K. Okamoto, M. Goldshtein, and P. Tsiotras, "Optimal covariance control for stochastic systems under chance constraints," *IEEE Control Systems Letters*, vol. 2, no. 2, pp. 266–271, 2018.

[88] K. Okamoto and P. Tsiotras, "Optimal stochastic vehicle path planning using covariance steering," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2276–2281, 2019.

[89] M. Pachter and K. D. Pham, "Discrete-time linear-quadratic dynamic games," *Journal of Optimization Theory and Applications*, vol. 146, no. 1, pp. 151–179, 2010.

[90] A. F. Hotz and R. E. Skelton, "A covariance control theory," in *IEEE Conference on Decision and Control*, vol. 24, Fort Lauderdale, FL, 1985, pp. 552–557.

[91] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal steering of a linear stochastic system to a final probability distribution, Part I," *IEEE Transactions on Automatic Control*, vol. 61, no. 5, pp. 1158–1169, 2016.

[92] ——, "Optimal steering of a linear stochastic system to a final probability distribution, Part II," *IEEE Transactions on Automatic Control*, vol. 61, no. 5, pp. 1170–1180, 2016.

[93] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal steering of a linear stochastic system to a final probability distribution, Part III," *IEEE Transactions on Automatic Control*, vol. 63, no. 9, pp. 3112–3118, 2018.

[94] J. Ridderhof, K. Okamoto, and P. Tsiotras, "Nonlinear uncertainty control with iterative covariance steering," in *IEEE Conference on Decision and Control*, Nice, France, 2019, pp. 3484–3490.

[95] J. Ridderhof and P. Tsiotras, "Minimum-fuel powered descent in the presence of random disturbances," in *AIAA Scitech Forum*, AIAA Paper 2019-0646, San Diego, CA, 2019.

[96] J. Ridderhof, J. Pilipovsky, and P. Tsiotras, "Chance-constrained covariance control for low-thrust minimum-fuel trajectory optimization," in *2020 AAS/AIAA Astrodynamics Specialist Conference*, 2020.

[97] J. Ridderhof and P. Tsiotras, "Minimum-fuel closed-loop powered descent guidance with stochastically derived throttle margins," *Journal of Guidance, Control, and Dynamics*, vol. 44, no. 3, pp. 537–547, 2020.

[98] A. Halder and E. D. Wendel, "Finite horizon linear quadratic Gaussian density regulator with Wasserstein terminal cost," in *American Control Conference*, IEEE, Boston, MA, 2016, pp. 7249–7254.

[99] M. Goldshtein and P. Tsiotras, "Finite-horizon covariance control of linear time-varying systems," in *IEEE Conference on Decision and Control*, Melbourne, Australia, 2017, pp. 3606–3611.

[100] E. Bakolas, "Finite-horizon covariance control for discrete-time stochastic linear systems subject to input constraints," *Automatica*, vol. 91, pp. 61–68, 2018.

[101] P. E. Moseley, *The Apollo entry guidance: A review of the mathematical development and its operational characteristics*, TRW Note No. 69-FMT-791, Dec. 1969.

[102] G. L. Carman, D. G. Ives, and D. K. Geller, "Apollo-derived Mars precision lander guidance," in *Atmospheric Flight Mechanics Conference*, 1998-4570, Boston, MA, 1998.

[103] G. F. Mendeck and L. Craig McGrew, "Entry guidance design and postflight performance for 2011 Mars Science Laboratory mission," *Journal of Spacecraft and Rockets*, vol. 51, no. 4, pp. 1094–1105, 2014.

[104] C. Winsor and R. Roy, "The application of specific optimal control to the design of desensitized model following control systems," *IEEE Transactions on Automatic Control*, vol. 15, no. 3, pp. 326–333, 1970.

[105] R. Subbayyan, V. V. S. Sarma, and M. C. Vaithilingam, "An approach for sensitivity-reduced design of linear regulators," *International Journal of Systems Science*, vol. 9, no. 1, pp. 65–74, 1978.

[106] C. Verde and P. M. Frank, "Sensitivity reduction of the linear quadratic regulator by matrix modification," *International Journal of Control*, vol. 48, no. 1, pp. 211–223, 1988.

[107] G. Zames, "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses," *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 301–320, 1981.

[108] B.-C. Chang and J. Pearson, "Optimal disturbance reduction in linear multivariable systems," *IEEE Transactions on Automatic Control*, vol. 29, no. 10, pp. 880–887, 1984.

[109] D. L. Tang, "The design of a desensitized controller - A sensitivity function approach," in *American Control Conference*, San Diego, CA, 1990, pp. 3103–3111.

[110] F. Cheli, F. Resta, and E. Sabbioni, "Observer design for active suspension control for passenger cars," in *ASME Conference on Engineering System Design and Analysis*, Torino, Italy, 2006, pp. 1–9.

[111] H. Seywald and R. R. Kumar, "Desensitized optimal trajectories," *Advances in the Astronautical Sciences*, vol. 93, pp. 103–115, 1996.

[112] K. Seywald and H. Seywald, "Desensitized optimal control," in *AIAA Scitech Forum*, AIAA Paper 2019-0651, San Diego, CA, 2019.

[113] H. Seywald, "Desensitized optimal trajectories with control constraints," *Advances in the Astronautical Sciences*, vol. 114, pp. 737–743, 2003.

[114] H. Shen, H. Seywald, and R. W. Powell, "Desensitizing the minimum-fuel powered descent for mars pinpoint landing," *Journal of Guidance, Control, and Dynamics*, vol. 33, no. 1, pp. 108–115, 2010.

[115] H. Xu and H. Cui, "Robust trajectory design scheme under uncertainties and perturbations for Mars entry vehicle," in *IEEE International Conference on Computational Intelligence Communication Technology*, Ghaziabad, India, 2015, pp. 762–766.

[116] S. Li and Y. Peng, "Mars entry trajectory optimization using DOC and DCNLP," *Advances in Space Research*, vol. 47, no. 3, pp. 440–452, 2011.

[117] H. Ohlsson, "Regularization for sparseness and smoothness - applications in system identification and signal processing," Ph.D. dissertation, Department of Electrical Engineering, Linköping University, Sweden, 2010.

[118] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT press Cambridge, 2000.

[119] G. S. Aoude, B. D. Luders, J. M. Joseph, N. Roy, and J. P. How, "Probabilistically safe motion planning to avoid dynamic obstacles with uncertain motion patterns," *Autonomous Robots*, vol. 35, no. 1, pp. 51–76, 2013.

[120] W. Xu, J. Pan, J. Wei, and J. M. Dolan, "Motion planning under uncertainty for on-road autonomous driving," in *IEEE International Conference on Robotics and Automation*, Hong Kong, China, 2014, pp. 2507–2512.

[121] 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*, vol. 30, no. 7, pp. 895–913, 2011.

[122] A. K. Akametalu, J. F. Fisac, J. H. Gillula, S. Kaynama, M. N. Zeilinger, and C. J. Tomlin, "Reachability-based safe learning with gaussian processes," in *IEEE Conference on Decision and Control*, Los Angeles, CA, 2014, pp. 1424–1431.

[123] Y. Luo, P. Cai, A. Bera, D. Hsu, W. S. Lee, and D. Manocha, "PORCA: Modeling and planning for autonomous driving among many pedestrians," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3418–3425, Oct. 2018.

[124] Q. Zhu, "Hidden Markov model for dynamic obstacle avoidance of mobile robot navigation," *IEEE Transactions on Robotics and Automation*, vol. 7, no. 3, pp. 390–397, Jun. 1991.

[125] C. Fulgenzi, A. Spalanzani, and C. Laugier, "Dynamic obstacle avoidance in uncertain environment combining PVOs and occupancy grid," in *IEEE International Conference on Robotics and Automation*, Roma, Italy, 2007, pp. 1610–1616.

[126] G. S. Aoude, B. D. Luders, D. S. Levine, and J. P. How, "Threat-aware path planning in uncertain urban environments," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipei, Taiwan, 2010, pp. 6058–6063.

[127] J. van den Berg, D. Ferguson, and J. Kuffner, "Anytime path planning and replanning in dynamic environments," in *IEEE International Conference on Robotics and Automation*, Orlando, FL, 2006, pp. 2366–2371.

[128] M. Zucker, J. Kuffner, and M. Branicky, "Multipartite RRTs for rapid replanning in dynamic environments," in *IEEE International Conference on Robotics and Automation*, Roma, Italy, 2007, pp. 1603–1609.

[129] N. E. Du Toit and J. W. Burdick, "Robotic motion planning in dynamic, cluttered, uncertain environments," in *IEEE International Conference on Robotics and Automation*, Anchorage, AK, 2010, pp. 966–973.

[130] D. Fridovich-Keil, J. F. Fisac, and C. J. Tomlin, "Safely probabilistically complete real-time planning and exploration in unknown environments," in *IEEE International Conference on Robotics and Automation*, Montreal, QC, Canada, 2019, pp. 7470–7476.

[131] M. C. Mora and J. Tornero, "Path planning and trajectory generation using multirate predictive artificial potential fields," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nice, France, 2008, pp. 2990–2995.

[132] M. Rafieisakhaei, S. Chakravorty, and P. Kumar, "Non-Gaussian SLAP: Simultaneous localization and planning under non-Gaussian uncertainty in static and dynamic environments," *arXiv preprint arXiv:1605.01776*, 2016.

[133] J. C. Harsanyi, "Games with incomplete information played by "Bayesian" players, Parts I–III," *Management science*, vol. 14, no. 3, 5, 7, pp. 159–182, 320–334, 486–502, 1967-1968.

[134] S. Zamir, "Bayesian games: Games with incomplete information," in *Encyclopedia of Complexity and Systems Science*, R. A. Meyers, Ed. New York, NY: Springer New York, 2009, pp. 1–26.

[135] D. T. Larsson, G. Kotsalis, and P. Tsiotras, "Nash and correlated equilibria for pursuit-evasion games under lack of common knowledge," in *IEEE Conference on Decision and Control*, Miami Beach, FL, 2018, pp. 3579–3584.

[136] A. Nayyar, A. Gupta, C. Langbort, and T. Başar, "Common information based Markov perfect equilibria for stochastic games with asymmetric information: Finite games," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 555–570, 2013.

[137] A. Gupta, A. Nayyar, C. Langbort, and T. Başar, "Common information based Markov perfect equilibria for linear-Gaussian games with asymmetric information," *SIAM Journal on Control and Optimization*, vol. 52, no. 5, pp. 3228–3260, 2014.

[138] A. Gupta, C. Langbort, and T. Başar, "Dynamic games with asymmetric information and resource constrained players with applications to security of cyberphysical

systems," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 71–81, 2016.

[139] D. Vasal, A. Sinha, and A. Anastasopoulos, "A systematic process for evaluating structured perfect Bayesian equilibria in dynamic games with asymmetric information," *IEEE Transactions on Automatic Control*, vol. 64, no. 1, pp. 81–96, 2018.

[140] Shuzhi Sam Ge, Hongbin Ma, and Kai-Yew Lum, "Detectability in games of pursuit evasion with antagonizing players," in *IEEE Conference on Decision and Control*, New Orleans, LA, 2007, pp. 1404–1409.

[141] S. Bhattacharya and S. Hutchinson, "A cell decomposition approach to visibility-based pursuit evasion among obstacles," *The International Journal of Robotics Research*, vol. 30, no. 14, pp. 1709–1727, 2011.

[142] K. Mizukami and V. Tews, "State-estimation in a pursuit-evasion-game with incomplete information-exchange," in *Optimization Techniques*, Springer, 1980, pp. 241–250.

[143] D. Shen, B. Jia, G. Chen, E. Blasch, and K. Pham, "Pursuit-evasion games with information uncertainties for elusive orbital maneuver and space object tracking," in *Sensors and Systems for Space Applications VIII*, K. D. Pham and G. Chen, Eds., International Society for Optics and Photonics, vol. 9469, SPIE, 2015, pp. 151–160.

[144] T. D. Woodbury and J. E. Hurtado, "Adaptive play via estimation in uncertain nonzero-sum orbital pursuit evasion games," in *AIAA SPACE and Astronautics Forum and Exposition*, AIAA Paper 2017-5247, Orlando, FL, 2017.

[145] D. Linville and J. Hess, "Linear regression models applied to spacecraft imperfect information pursuit-evasion differential games," in *AIAA Scitech Forum*, AIAA Paper 2020-0952, Orlando, FL, 2020.

[146] D. Ye, X. Tang, Z. Sun, and C. Wang, "Multiple model adaptive intercept strategy of spacecraft for an incomplete-information game," *Acta Astronautica*, vol. 180, pp. 340–349, 2021.

[147] J. Shinar and V. Y. Glizer, "Solution of a delayed information linear pursuit-evasion game with bounded controls," *International Game Theory Review*, vol. 1, no. 03n04, pp. 197–217, 1999.

[148] K. A. Cavalieri, N. Satak, and J. E. Hurtado, "Incomplete information pursuit-evasion games with uncertain relative dynamics," in *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2014-0971, National Harbor, Maryland, 2014.

[149] G. Hexner, I. Rusnak, and H. Weiss, "A pursuit-evasion game with incomplete information," in *Mediterranean Conference on Control and Automation*, Akko, Israel, 2019, pp. 583–588.

[150] O. Gurel-Gurevich, "Pursuit–evasion games with incomplete information in discrete time," *International Journal of Game Theory*, vol. 38, no. 3, pp. 367–376, 2009.

[151] X. Huang, P. Maupin, and R. van der Meyden, "Model checking knowledge in pursuit evasion games," in *International Joint Conference on Artificial Intelligence*, Barcelona, Catalonia, Spain, 2011, pp. 240–245.

[152] L. S. Shapley, "Stochastic games," *Proceedings of the National Academy of Sciences*, vol. 39, no. 10, pp. 1095–1100, 1953.

[153] E. Altman and E. Solan, "Constrained games: The impact of the attitude to adversary's constraints," *IEEE Transactions on Automatic Control*, vol. 54, no. 10, pp. 2435–2440, 2009.

[154] D. W. Repperger and A. J. Koivo, "Optimal terminal rendezvous as a stochastic differential game problem," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 8, no. 3, pp. 319–326, 1972.

[155] J.-S. Ha, H.-J. Chae, and H.-L. Choi, "A stochastic game-based approach for multiple beyond-visual-range air combat," *Unmanned Systems*, vol. 6, no. 1, pp. 67–79, 2018.

[156] B. A. Free, M. J. McHenry, and D. A. Paley, "Non-deterministic predator-prey model with accelerating prey," in *American Control Conference*, IEEE, Milwaukee, WI, 2018, pp. 1202–1207.

[157] S. M. Perlaza, H. Tembine, S. Lasaulce, and M. Debbah, "Quality-of-service provisioning in decentralized networks: A satisfaction equilibrium approach," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 2, pp. 104–116, 2011.

[158] R. M. Ford and R. N. Strickland, "Representing and visualizing fluid flow images and velocimetry data by nonlinear dynamical systems," *Graphical Models and Image Processing*, vol. 57, no. 6, pp. 462–482, 1995.

[159] C. Zhang and A. Hammad, "Improving lifting motion planning and re-planning of cranes with consideration for safety and efficiency," *Advanced Engineering Informatics*, vol. 26, no. 2, pp. 396–410, 2012.

[160] K. Hauser, "On responsiveness, safety, and completeness in real-time motion planning," *Autonomous Robots*, vol. 32, no. 1, pp. 35–48, 2012.

[161] S. Lengagne, N. Ramdani, and P. Fraisse, "Guaranteed computation of constraints for safe path planning," in *IEEE-RAS International Conference on Humanoid Robots*, Pittsburgh, PA, 2007, pp. 312–317.

[162] B. D. Luders, S. Karaman, and J. P. How, "Robust sampling-based motion planning with asymptotic optimality guarantees," in *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2013-5097, Boston, MA, 2013.

[163] J. S. Mitchell, "Geometric shortest paths and network optimization," *Handbook of Computational Geometry*, vol. 334, pp. 633–702, 2000.

[164] H. K. Khalil, *Nonlinear Systems*. Prentice Hall, 2002, Chapter 3.

[165] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control*. Taylor & Francis, 1975.

[166] D. W. Peterson, "On sensitivity in optimal control problems," *Journal of Optimization Theory and Applications*, vol. 13, no. 1, pp. 56–73, 1974.

[167] N. A. Shneydor, *Missile Guidance and Pursuit: Kinematics, Dynamics and Control*. Horwood Publishing Limited, 1998, Chapters 3, 4.

[168] E. Bakolas and P. Tsiotras, "Feedback navigation in an uncertain flowfield and connections with pursuit strategies," *Journal of Guidance, Control and Dynamics*, vol. 35, no. 4, pp. 1268–1279, 2012.

[169] W. Sun and P. Tsiotras, "Sequential pursuit of multiple targets under external disturbances via zermelo–voronoi diagrams," *Automatica*, vol. 81, pp. 253–260, 2017.

[170] M. A. Patterson and A. V. Rao, "GPOPS-II: A MATLAB software for solving multiple-phase optimal control problems using hp-adaptive Gaussian quadrature collocation methods and sparse nonlinear programming," *ACM Transactions on Mathematical Software*, vol. 41, no. 1, 1:1–1:37, 2014.

[171] L. Cesari, *Optimization-Theory and Applications: Problems with Ordinary Differential Equations*. NY: Springer-Verlag, 1983, Chapter 9.

[172] V. R. Makkapati, W. Sun, and P. Tsiotras, "Pursuit-evasion problems involving two pursuers and one evader," in *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2018-2107, Kissimmee, FL, 2018.

[173] V. A. Toponogov, *Differential Geometry of Curves and Surfaces: A Concise Guide*, V. Y. Rovenski, Ed. Boston: Birkhäuser, 2006, Chapter 1.

[174] E. W. Weisstein, *Circle-circle intersection*, From MathWorld–A Wolfram Web Resource.

[175] R. W. Morgan and J. L. Riel, "Blind evasion by random switching maneuvers," in *American Control Conference*, IEEE, Boston, MA, 2016, pp. 3126–3131.

[176] W. H. Fleming and P. E. Souganidis, "On the existence of value functions of two-player, zero-sum stochastic differential games," *Indiana University Mathematics Journal*, vol. 38, no. 2, pp. 293–314, 1989.

[177] T. Başar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1998.

[178] R. Behn and Y.-C. Ho, "On a class of linear stochastic differential games," *IEEE Transactions on Automatic Control*, vol. 13, no. 3, pp. 227–240, 1968.

[179] J. R. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics*. John Wiley & Sons, 2019.

[180] S. Li and T. Başar, "Distributed algorithms for the computation of noncooperative equilibria," *Automatica*, vol. 23, no. 4, pp. 523–533, 1987.

[181] J. Löfberg, "YALMIP : A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[182] MOSEK, ApS. (2017). "The MOSEK optimization toolbox for MATLAB manual. version 8.1."

[183] V. Turetsky and J. Shinar, "Missile guidance laws based on pursuit–evasion game formulations," *Automatica*, vol. 39, no. 4, pp. 607–618, 2003.

[184] S. Osher, R. Fedkiw, and K. Piechor, "Level set methods and dynamic implicit surfaces," *Appl. Mech. Rev.*, vol. 57, no. 3, B15–B15, 2004.

[185] T. Lolla, M. P. Ueckermann, K. Yiğit, P. J. Haley, and P. F. Lermusiaux, "Path planning in time dependent flow fields using level set methods," in *IEEE International Conference on Robotics and Automation*, Piscataway, NJ, 2012, pp. 166–173.

[186] T. Corpetti, E. Mémin, and P. Pérez, "Extraction of singular points from dense motion fields: An analytic approach," *Journal of Mathematical Imaging and Vision*, vol. 19, no. 3, pp. 175–198, 2003.

[187] S. Beirle, J. Lampel, C. Lerot, H. Sihler, and T. Wagner, "Parameterizing the instrumental spectral response function and its changes by a super-Gaussian and its derivatives," *Atmospheric Measurement Techniques*, vol. 10, no. 2, pp. 581–598, 2017.

[188] L. S. Pontryagin, *Mathematical Theory of Optimal Processes*. Routledge, 2018.

[189] S. Shivam, A. Kanellopoulos, K. G. Vamvoudakis, and Y. Wardi, "A predictive deep learning approach to output regulation: The case of collaborative pursuit evasion," in *IEEE Conference on Decision and Control*, Nice, France, 2019, pp. 853–859.

[190] N. T. Kokolakis, A. Kanellopoulos, and K. G. Vamvoudakis, "Bounded rational unmanned aerial vehicle coordination for adversarial target tracking," in *American Control Conference*, IEEE, Denver, CO, 2020, pp. 2508–2513.

[191] M. Kothari, J. G. Manathara, and I. Postlethwaite, "A cooperative pursuit-evasion game for non-holonomic systems," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 1977–1984, 2014, 19th IFAC World Congress.

[192] N. Grigorenko, Y. N. Kiselev, N. Lagunova, D. Silin, and N. Trin'ko, "Solution methods for two-player differential games," *Computational Mathematics and Modeling*, vol. 8, no. 1, pp. 34–48, 1997.

[193] Ş. MiricĂ, "Reducing a differential game to a pair of optimal control problems," in *Differential Equations, Chaos and Variational Problems*, V. Staicu, Ed., Basel: Birkhäuser Basel, 2008, pp. 269–283.

[194] G. Ibragimov, M. Ferrara, M. Ruziboev, and B. A. Pansera, "Linear evasion differential game of one evader and several pursuers with integral constraints," *International Journal of Game Theory*, pp. 1–22, 2021.

[195] V. R. Makkapati, J. Ridderhof, P. Tsiotras, J. Hart, and B. van Bloemen Waanders, "Desensitized trajectory optimization for hypersonic vehicles," in *IEEE Aerospace Conference*, Big Sky, MT, 2021, pp. 1–10.

VITA

Venkata Ramana Makkapati is a doctoral candidate in the School of Aerospace Engineering at Georgia Tech. He applies techniques from optimal control and differential games to address problems pertaining to autonomous vehicles, emphasizing safe path planning and airspace security. He received the B.Tech. from IIT Madras in 2014, and the M.Tech. from IIT Kanpur in 2016, both in Aerospace Engineering. In the past, he held internship positions at Foresight AI (a self-driving car technology start-up in Silicon Valley), CSIR - National Aerospace Laboratories, Bangalore, India, and Mahindra & Mahindra, Chennai, India. Ramana is an FAA certified Private Pilot in the category Airplane Single Engine Land.