Receding Horizon based Cooperative Vehicle Control with Optimal Task Allocation

Mohammad Khosravi

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Applied Science at

Concordia University

Montréal, Québec, Canada

March 2016

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Mohammad Khosravi

Entitled: Receding Horizon based Cooperative Vehicle Control with Optimal Task Allocation

and submitted in partial fulfilment of the requirements for the degree of

Master of Applied Science

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

____________________ Dr. R. Raut, Chair

____________________ Dr. W. F. Xie, External Examiner

____________________ Dr. K. Skonieczny, Examiner

____________________ Dr. A. G. Aghdam, Supervisor

Approved by: ____________________

Dr. W. E. Lynch, Chair

Department of Electrical and Computer Engineering

________ 20 ____________________

Dr. A. Asif

Dean, Faculty of Engineering and Computer Science

ABSTRACT

Receding Horizon based Cooperative Vehicle Control with Optimal Task Allocation

Mohammad Khosravi

The problem of cooperative multi-target interception in an uncertain environment is investigated in this thesis. The targets arrive in the mission space sequentially at *a priori* unknown time instants and *a priori* unknown locations, and then move on *a priori* unknown trajectories. A group of vehicles with known dynamics are employed to visit the targets as quickly and efficiently as possible. To this end, a time-discounting reward is defined for each target which can be collected only if one of the vehicles visits that target. A cooperative receding horizon scheme is designed, which predicts the future positions of the targets and maximizes the estimate of the expected total collectible rewards, accordingly. The problem is initially investigated for the case when there are a finite number of targets arriving in the mission space sequentially. It is shown that the number of targets that are not visited by any vehicle in the mission space will be sufficiently small if the targets arrive sufficiently infrequently. The problem is then generalized to the case of infinite number of targets and a finite-time convergence analysis is also presented. A more practical case where the vehicles have limited sensing and communication ranges is also investigated using a game-theoretic approach. The problem is then solved for the case when a cluster of vehicles is required to visit each target. Simulations confirm the efficacy of the proposed strategies.

To Sara,

Mom & Dad

ACKNOWLEDGEMENTS

My appreciation, first and foremost goes to my knowledgeable, erudite and exceptional supervisor, Prof. Amir G. Aghdam, for his kind supports and patient guidances. His encouragements, motivation and immense knowledge, were the invaluable assets I had throughout my experience in his research group and I am deeply grateful for them.

I would like to thank Dr. Hassan Rivaz for his insightful suggestions and the great helpful discussions. Also, I do appreciate my colleagues and friends at Concordia University, Dr. Mohammad Mehdi Asadi and Mr. Hossein Khodadadi, for their patience and kindness in our collaborations. I was very fortunate to have the most rewarding discussions with them.

I am grateful to my friends, Hossein, Mohsen, Meysam and Roozbeh, with whom I had great moments in Montréal and specifically Hossein who was also like a brother to me.

I ultimately would like to thank my beloved parents, brother and sisters for their love and warm supports.

Lastly, I like to thank Sara, my true love, for her presence and supports in every moment of my life.

I would like to acknowledge the financial support from Natural Sciences and Engineering Research Council of Canada (NSERC) and Defence Research & Development Canada (DRDC) throughout my graduate work at Concordia University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

The last two decades have witnessed an ever-increasing interest in multi-agent systems [1–12] inspired by and applied to a wide variety of fields of science and technology such as biology [13–17], control [18–22], robotics [23–27], computer science [28,29], economy, marketing and finance [30,31]. The main goal of multi-agent systems is to achieve a global objective with a set of simple and limited components and the proper use of information exchange between the agents. Multi-agent systems and methods are explored for a variety of applications related to control design problems such as surveillance [32, 33], search and rescue [34–36], reconnaissance missions [37, 38], sensor networks [39, 40], automated highway systems [41], environmental sampling [42, 43], motion coordination of robots [44–46], formation control of satellite clusters [47, 48], air traffic control [49], consensus [50], network connectivity control [51], target assignment and cooperative multi-target interception [52–54], to name only a few.

The multi-target interception problem with multiple vehicles is an emerging topic in the cooperative control literature. The problem is concerned with a group of vehicles, which are desired to cooperatively visit some targets that appear in the mission space at random time instants.

The *pursuit-evasion* problem is a well-known and vastly studied topic in the contexts of computer science, mathematics, artificial intelligence, robotics, control, physics, etc. [55–59]. In the literature, the problem has been investigated under different titles such as *cop and robber* [60], *lion and man* [61, 62], *graph searching* [63], *lady and bandit* [64, 65],and *chases and escapes* [57]. Usually, this type of problem is analyzed using multifarious formulations [55–57]. In all variants of the pursuit-evasion problem, a group of pursuers attempting to track down a group of evaders [55]. In addition to pursuer(s) and evaders(s), a set of one or more defenders like missiles or bodyguards may also be considered which are supposed to defend the evader(s) against the attacker(s) [66]. The environment, on the other hand, can be a discrete space like a graph [67] or a continuous space like a manifold [55]. Similarly, the pursuit and evasion procedure can be discrete-time [68] or continuous-time [69]. The information of the pursuer(s), evader(s) and possible defender(s) on each other as well as the environment is not necessarily perfect [70, 71]. The motions of targets are sometimes independent of the strategies of pursuers, while some other times, the pursuers and evaders may have conflicting objectives [72]. Formulations of pursuit-evasion problems may differ by the constraints and the level of maneuverability considered for the pursuers and evaders [73, 74]. In particular the main focus of this thesis is directed towards the multi-pursuers and multi-evaders problem with perfect information knowledge, continuous-time procedure and continuous space, where evaders moves possibly independently of pursuers. In [74] a practical ad-hoc pursuit algorithm is introduced for the pursuers to capture a finite number of evaders and super-evaders (the evaders with greater velocities compared to the pursuers). It is to be noted the pursuit-evasion procedure in [74] is in the discrete-time framework, and also the magnitude of the velocity of each pursuer and each evader is assumed to be constant. Moreover, no justification or theoretical proof is provided in [74] for the optimality or sub-optimality of the algorithm. In [75], a hierarchical approach is introduced to tackle the problem using the combinatorial optimization. In order to hierarchically decompose the problem and obtain a suboptimal engagement between the pursuers and evaders, the authors in [72] derive a combinatorial optimization problem. Most of the results in the literature on pursuit-evasion games suffer from the curse of dimensionality, significant computational loads and lack of practical on-line implementations [72, 75]. Moreover, in order to define the game in pursuit-evasion problems, it is assumed that the evaders are intelligent and rational, which implies that their behaviors and policies are known [72]; this is not the case, however, in many applications.

The target interception problem has also been investigated in the operations research and computer science. One of the most widely studied related problems in this area is the famous Traveling Salesman Problem (TSP), where it is desired to find the shortest tour passing through a number of cities assuming that the list of cities and their mutual distances are given [76]. It is shown in [77] that TSP is an NP-complete combinatorial optimization problem. Different formulations are proposed for the problem, including one in the context of integer programming optimization [78]. The multifarious variations of the TSP are extensively studied in the literature. In the Asymmetric Traveling Salesman Problem (ATSP), the distance between any pair of cities depends on the traveling direction [79]. In the formulation of time-constrained TSP, a time window is considered for any city in which the corresponding city is to be visited [80]. In another variant of problem, a profit is associated with visiting each city and the objective is to find a subset of cities for simultaneously maximizing the collected profits and minimizing the travel costs [81]. Similarly, in the orienteering problem, a reward is considered for each city and it is desired to determine a length-limited path for visiting a set of cities and collecting the corresponding rewards such that the total collected rewards is maximized [82]. The multiple TSP (mTSP) is an extension of the problem with more than one salesman [83]. The Vehicle Routing Problem (VRP) is another extension where there are a number of visiting points, referred to as way-points, and a number of vehicles in a depot; the vehicles are to visit the way-points in an optimal fashion [84, 85]. The vehicle routing problem has many variants including VRP with Pickup and Delivery (VRPPD) [86], VRP with Time Windows (VRPTW) [87], Capacitated VRP (CVRP) [88], VRP with Multiple Trips (VRPMT) [89], Open VRP (OVRP) [90] and Dynamic VRP (DVRP) [91]. In all of these problems, the target points are located inside the space from the beginning of the operation. In the m-vehicle Dynamic Traveling Repairman Problem (m-DTRP), on the other hand, is one of the most general forms of VRP, where a number of vehicles travel with bounded velocity in a bounded environment. The vehicles are supposed to service a set of demands with stochastic arrival times and stochastic location [92–97]. For m-DTRP, adaptive and receding-horizon-based policies are introduced in [97] and their optimality is proved for the cases of light and heavy demand loads. In addition to the above papers where only static points are considered, in [98] a different type of traveling salesman problem is introduced where each target point moves with a constant velocity. The Moving-Target VRP is studied in [99], where each target appears on a line and then moves in the space with a constant velocity. A variant of TSP is discussed in [100], where the targets move with the same constant velocity, and then a robotic arm moves in the space to collect the targets and deliver them to a certain depot. A similar case is discussed in [101], where the arm is assumed to have a limited capacity. The problem of dynamic vehicle routing with moving targets is investigated in [102], where each target should be visited in a certain time window. In [103, 104], DVRP with moving targets is investigated, where each target appears on a disc according to a spatiotemporal probability distribution and moves radially with constant speed toward escaping the disc while a single vehicle aims at capturing them before they escape. In all of the above papers, it was assumed that either the target points are fixed, or if they are moving, their velocities are known and fixed.

Reward collection in multi-agent systems provides a framework for a variety of problems such as coverage, data collection and multi-target interception [105–109]. In operations research, on the other hand, the prize-collecting TSP [110] and orienteering problem [82] are addressed. In this framework, one or more agents are collect rewards by accomplishing a number of tasks. The reward of each task can be fixed or it can depend on some other parameters and variables such as time or location [107–109,111,112]. They may also be some constraints in this type of problems to introduce a feasible space for collecting rewards. For example, in [105] it is assumed that some obstacles of polygonal shape exist, imposing physical constraints on the motions of vehicles, and consequently on reward collection. In the target interception problem, the rewards can be properly associated with the targets. In [108], the cooperative multi-target interception problem is tackled, with no *a priori* knowledge about the arrival times of the target points, using a cooperative receding horizon (CRH) control scheme. In [113], the authors further improve the reward-collection-based controller developed in [108], overcoming some of its limitations such as poor performance and instability in target trajectories.

In addition to the reward-collection frameworks discussed above, there are some other approaches [114]. In [115], the moving objects with known kinematics are assigned to the vehicles using dynamic Voronoi partitioning. The multi-target interception problem for a set of homogeneous moving targets with unicycle model is addressed in [116] by means of a distributed cooperative strategy. In [117], a Mixed-Integer Linear Programing (MILP) approach is used to find trajectories for a group of vehicles such that they visit a set of fixed way-points with some time constraints. Similar to [117], a MILP formulation is also used in [118] to provide a cooperative control approach for guarding a defense zone from a group of invaders.

In all of the above-mentioned papers, some restrictive assumptions are made: the targets are assumed to be stationary points in [92, 108]; the targets move with constant velocity [99, 103, 104]; only a single agent is to accomplish the mission in [103, 104]; the arrival times of the targets are assumed to be known in [115, 116], and certain conditions are imposed on targets' dynamics [72, 74]. Moreover, no performance metric is considered for the target-vehicle assignments in [74], and the designed algorithms in [117, 119] are computationally demanding. Considering these restrictions and drawbacks, it is desired to utilize a time-decomposition-based method, such as a receding horizon scheme, for designing a controller for an uncertain multi-target interception systems where the targets arrive in the mission space sequentially at *a priori* unknown time instants, in *a*

	Analytical results	No assumption on intelligence of targets	A priori unknown arrival times	A priori unknown arrival positions	A priori unknown trajectories	Unknown dynamics	Computationally tractable	Multi-target and Multi-vehicle
[75], [72]	✗	✗	✗	✗	✓	✓	✗	✓
[74]	✗	✓	✗	✗	✓	✓	✗	✓
[92], [93], [94], [95], [96], [97]	✓	✓	✓	✓	✗	✗	✓	✓
[103, 104]	✓	✓	✓	✓	✗	✗	✓	✗
[116]	✓	✓	✗	✗	✓	✗	✓	✓
[115]	✓	✗	✗	✗	✓	✗	✓	✓
[107], [113], [108], [109]	✓	✓	✓	✗	✗	✗	✓	✓

Table 1.1: Comparison Table

priori unknown position, and then move on *a priori* unknown trajectories. A comparison between the main characteristics of the existing results discussed above is summarized in Table 1.1.

1.1 Outlines of Thesis

In Chapter 2, a cooperative receding horizon controller (CRHC) is designed to track moving targets with unknown dynamics using a team of vehicles. Each target is assigned a time decreasing reward, which is collectible only if the target is visited by some vehicles, and the team objective is to maximize the total collected rewards. At each iteration, the vehicles face multiple targets, some of which may be new in the target space. Each target has an *a priori* unknown trajectory with a bounded velocity. As the targets may arrive sequentially in time, vehicles should visit them in minimal time to avoid a burst of unvisited target population and at the same time to have a stationary state.

In Chapter 3, a Receding Horizon-based Dynamic Decision-making Controller (RHDDC) is designed for heading control of a single vehicle toward intercepting targets which arrive in the mission space sequentially, moving with unknown dynamics. Similar to [103], a single vehicle is used to capture the sequence of targets with arrival times modeled stochastically. The mission space, on the other hand, is assumed to be a compact set in an Euclidean space (as opposed to a disk). The arrival times of the targets are modeled by a renewal process which is a generalization of the Poisson process. One of the important characteristics of the present problem setting is that no spatial distribution for the initial positions of targets is considered. Furthermore, target trajectories and dynamics are assumed to be *a priori* unknown. Similar to Chapter 2, the designed strategy is based on assigning rewards for capturing the targets, and predicting the future target positions. Convergence analysis is provided, and simulations for different scenarios concerning frequent and infrequent target arrivals are presented.

In Chapter 4, a Cooperative Receding Horizon Controller is designed for heading control of a set of vehicles toward intercepting targets, arriving the mission space in *a priori* unknown times and *a priori* unknown positions, and also moving with *a priori* unknown dynamics. Here, similar to Chapter 2, a team of vehicles are supposed to capture a set of targets moving with *a priori* unknown trajectories, and further generalize it by assuming that the arrival positions and times are *a priori* unknown. Moreover, vehicles have limited ranges for sensing the targets and also limited ranges for communication, i.e., each vehicle can only sense the targets located in a region around it and also communicate only with vehicles with distance less than a prescribed range. Dealing with this level of uncertainties in the environment and vehicles limitations, a distributed on-line controller using receding horizon is required. Toward this goal, the method introduced in Chapter 2 has been extended by exploiting recent developments in games theory. In this approach, each of the targets is assigned a time decreasing reward, which is collectible only if the target is visited by some vehicles, and considering these rewards and problem constraints, a utility function is designed with respect to each vehicle. The resulting structure forms a potential game with total collectible reward as its potential functions. Using appropriate learning dynamics, vehicles decide upon their strategies and consequently on their headings.

Next, in Chapter 5, the cooperative multi-target interception problem in uncertain environment with double-integrator vehicles is investigated. Similar to Chapter 2, the problem is reformulated as a maximum reward collection problem which maximize the expected reward collectible from the set of available targets in the mission space. The reward function is a time discounting function assigned to each target and can be collected only if the target is visited by a vehicle. However, since targets are assumed to be moving objects with a priori unknown arrival times and trajectories, the existing uncertainties in the environment render the one-shot optimization rather impractical. Therefore, a cooperative receding horizon controller is utilized toward maximizing the collected reward and based on the prediction of the future positions of targets with the given limited information.

In Chapter 6, a Cooperative Receding Horizon (CRH) controller is presented, where agents are dynamically clustered and assigned to the targets to collect the respective rewards. Similar to [120–122], the proposed controller sequentially solves an optimization problem with a payoff function and a set of constraints. The constraints are updated in each iteration using the existing limited information over a planning horizon. The payoff function accounts for the estimation of maximum total reward expected to be collected by the end of mission, the clustering and assignments strategies, uniform configurations of agents in vicinity of the targets and finally, the imperfection of clusters. In the designed scheme, the agents are not forced to be committed to fixed clusters or targets, which is desirable for the uncertain environments.

Finally, the contributions of thesis are reviewed and summarized in Chapter 7. Also, further research directions are referred and introduced in Chapter 7.

1.2 Publications

The results of this master thesis and author's master's work are published (or submitted for publication) in a number of journals and conference proceedings [120–126], as listed below.

- M. Khosravi and A. G. Aghdam, *"Cooperative Receding Horizon Control for Multi-Target Interception in Uncertain Environments"*, in Proceedings of the 53rd IEEE Conference on Decision and Control, 2014.

- M. Khosravi and A. G. Aghdam, *"Stability Analysis of Dynamic Decision-Making for Vehicle Heading Control"*, in Proceedings of the American Control Conference, 2015.

- M. Khosravi, H. Khodadadi, H. Rivaz, and A. G. Aghdam, *"Cooperative Control for Multi-Target Interception with Sensing and Communication Limitations: A Game Theoretic Approach"*, in Proceedings of the 54th IEEE Conference on Decision and Control, 2015.

- M. Khosravi, H. Khodadadi, H. Rivaz, and A. G. Aghdam, *"Maximum Reward Collection Problem: A Cooperative Receding Horizon Approach for Dynamic Clustering"*, in Proceedings of the 2015 Conference on Research in Adaptive and Convergent Systems, 2015, pp. 38-43 (invited paper).

- M. Khosravi, H. Khodadadi, H. Rivaz, and A. G. Aghdam, *"Cooperative Receding Horizon Control of Double Integrator Vehicles for Multi-Target Interception"*, in Proceedings of the American Control Conference, 2015.

- M. M. Asadi, M. Khosravi, A. G. Aghdam, and S. Blouin, *"Joint power optimization and connectivity control problem over underwater random sensor networks"*, in Proceedings of the American Control Conference, 2015.

- M. M. Asadi, M. Khosravi, A. G. Aghdam, and S. Blouin, *"Generalized Algebraic Connectivity for Asymmetric Networks"*, in Proceedings of the American Control Conference, 2016.

- M. M. Asadi, M. Khosravi, and A. G. Aghdam, *"Flocking in Multi-Agent Networks with Limited Fields of View"* (submitted to a journal).

- M. Khosravi, H. Khodadadi, H. Rivaz, and A. G. Aghdam, *"Cooperative Receding Horizon Control for Multi-Target Interception in Uncertain and Adversarial Environments"* (submitted to a conference).

- M. M. Asadi, M. Khosravi, A. G. Aghdam, and S. Blouin, *"A Subspace Consensus Approach for Distributed Connectivity Assessment of Asymmetric Networks"* (submitted to a conference).

- M. M. Asadi, M. Khosravi, A. G. Aghdam, *"Connectivity of Asymmetric Networks: An Algebraic Measure"* (submitted to a journal).

Chapter 2

Cooperative Receding Horizon Control for Multi-Target Interception in Uncertain Environments

In this chapter, the problem of cooperative dynamic vehicle routing for tracking a set of moving objects with *a priori* unknown trajectories and dynamics is investigated. The notion of "visiting a target" is defined to describe the tasks and a cooperative receding horizon controller is designed to address the problem. The design is based on the prediction of the future positions of targets with limited information, and a reward allocation strategy for accomplishing the defined tasks. A target tracking scenario is considered, where a sequence of targets arrive in the mission space. It is shown that the number of targets which are not visited by any vehicle will remain sufficiently small in time, if the arrival of the targets is sufficiently infrequent.

2.1 Problem Formulation

Consider a set of N moving targets and a set of M vehicles in a *mission space*, denoted by $\mathcal{M}$, which is a closed convex subset of $\mathbb{R}^d$. Let $I_{\mathcal{T}} = \{1, 2, 3, \ldots, N\}$ and $I_{\mathcal{V}} = \{1, 2, 3, \ldots, M\}$ be the index sets for targets and vehicles, respectively. Let also $\mathrm{x}_j(t) \in \mathbb{R}^d$ and $\mathrm{y}_i(t) \in \mathbb{R}^d$ be respectively the position vectors of vehicle j and target i at any given time $t \in [0, T]$, for any $j \in I_{\mathcal{V}}$ and $i \in I_{\mathcal{T}}$, where T is a finite final time horizon for the accomplishment of the mission.

The dynamics of the j^{th} vehicle for $j \in I_{\mathcal{A}}$ is given by

$$\dot{\mathrm{x}}_j(t) = \mathrm{u}_j(t) = V_j(t)\mathrm{d}_j(t), \quad \forall j \in I_{\mathcal{V}}, \tag{2.1}$$

where $\mathrm{d}_j(t) \in S^{d-1} = \{\mathrm{d} \in R^d; \|\mathrm{d}\| = 1\}$ is the control input for the direction of the velocity vector and $V_j(t) \in [0, V_j]$ is the control input for its magnitude, for any $j \in I_{\mathcal{V}}$.

The trajectory of each target is a C^1 curve in the mission space $\mathcal{M}$ which is assumed to satisfy the following two geometric conditions.

Assumption 1. *(Global Geometric Condition) If* $\mathrm{y}_i(\tau) \in \mathcal{M}$ *for some* $\tau \in [0, T]$ *and any* $i \in I_{\mathcal{T}}$*, then* $\mathrm{y}_i(t) \in \mathcal{M}$ *for all* $t \in [\tau, T]$*.*

The global geometric condition on targets' trajectories guarantees that once a target is detected in the mission space, it will remain inside it until the end of the mission. Not only is this property dependent on the targets' trajectories, it also depends on the geometry of the mission space. In the special case when $\mathcal{M} = \mathbb{R}^d$, then the global geometric condition is satisfied automatically.

Assumption 2. *(Local Geometric Condition) There exists non-negative scalars* v, B *such that for any* $i \in I_{\mathcal{T}}$ *and* $\tau \in [0, T]$*,*

$$\|\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(\tau)\| \leq v, \tag{2.2}$$

and

$$\sup_{t \in (\tau, T]} \|\alpha_i(t, \tau)\| \leq B, \tag{2.3}$$

where $\alpha_i(t, \tau)$ *is a* C^1 *function satisfying the following equality*

$$\mathrm{y}_i(t) = \mathrm{y}_i(\tau) + \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(\tau)(t - \tau) + \frac{1}{2}\alpha_i(t, \tau)(t - \tau)^2. \tag{2.4}$$

Assume that $\mathrm{y}_i(t)$ is a C^2 function, and that there exist non-negative scalars v, B

such that for any $i \in I_{\mathcal{T}}$ and $\tau \in [0, T]$, the following conditions hold:

$$\|\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(\tau)\| \leq v, \quad \|\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{y}_i(\tau)\| \leq B. \tag{2.5}$$

Then, from Taylor's theorem with mean-value form of the remainder [127], $\mathrm{y}_i(t)$ satisfies Assumption 2.

Assumption 3. *The position and velocity vectors are available at the beginning of each time horizon (i.e., at time instant τ in (2.4)).*

As a result of Assumption 3, one can estimate the positions of the targets at any future instant within the finite horizon. Let this estimate be denoted by $\hat{\mathrm{y}}_i(\cdot)$ for any $i \in I_{\mathcal{T}}$. Then

$$\hat{\mathrm{y}}_i(t) = \mathrm{y}_i(\tau) + \mathrm{v}_i(\tau)(t - \tau), \quad t \in [\tau, T]. \tag{2.6}$$

Definition 1. *Given a positive scalar $s_i, i \in I_{\mathcal{T}}$, the j^{th} vehicle is said to visit the i^{th} target at time t, if $\|\mathrm{x}_j(t) - \mathrm{y}_i(t)\| \leq s_i$.*

Remark 1. *The scalar s_i in Definition 1 is introduced mainly for practical considerations in relation to the size of the target. More precisely, while the dynamic equation of each target is implicitly expressed as a point mass, the scalar s_i is used to account for the size of the i^{th} target as a rigid body. For instance, if the i^{th} target has a spherical shape with* *radius r_1 and each vehicle also has a spherical shape with radius r_2, then $s_i := r_1 + r_2$.*

Corresponding to each target, a *task* is defined which is completed only if the target is visited by at least one vehicle.

2.2 Cooperative Receding Horizon Scheme

In order for the vehicles to track the targets, a time-decreasing reward is assigned to each task which can be collected only if the target is visited (i.e., the task is accomplished). The goal of the team is to maximize the collected rewards. The vehicles plan their paths iteratively, where at the beginning of each iteration they calculate their headings and the size of movements such that an estimation of the future collectible rewards is maximized.

Let R_i be the maximum reward considered for task i before any deprivation results due to the passage of time. Let also $\rho_i : [0, T] \to [0, 1]$ be a decreasing function of time representing the rate of reward loss over time. One can now form a function $R_i\rho_i$, called *reward function*, which satisfies the desired properties discussed earlier. There are different candidate functions for ρ_i which model scheduling and time priorities. In particular, consider the following discount function

$$\rho_i(t) = 1 - \frac{f_i}{T}t, \quad i \in T \tag{2.7}$$

where $f_i \in (0, 1]$ is a target-specific loss parameter which is chosen to reflect different cases of interest.

2.2.1 Cooperative Structure

Given the positions of the targets and vehicles in the mission space $\mathcal{M}$, it is desired to properly assign tasks to the vehicles. More precisely, the objective is to find a set of assignments, each one denoted by

$$a_{ij} : \mathcal{M}^M \times \mathcal{M}^N \rightarrow [0, 1], \quad \forall i \in I_{\mathcal{T}}, \forall j \in I_{\mathcal{V}} \tag{2.8}$$

reflecting the amount of interest of vehicle j in target i being assigned to it, for any $i \in I_{\mathcal{T}}, j \in I_{\mathcal{V}}$.

There are a variety ofvisited hods for designing the function in (2.8) . For instance, one can use a Voronoi-based assignment, where each vehicle is typically assigned to one of its nearest targets. In this case, a map $\pi : I_{\mathcal{V}} \rightarrow I_{\mathcal{T}}$ given by

$$\pi(j) \in \operatorname*{argmin}_{i \in I_{\mathcal{T}}} \|\mathrm{x}_j - \mathrm{y}_i\|, \quad \forall j \in I_{\mathcal{V}},$$

one can define $a_{ij} = \delta_{\pi(j)j}$, where δ is the Kronecker delta function. The competition-based assignment, on the other hand, considers two nearest vehicles whose index belongs to $\mathcal{B}(\mathrm{y}_i) \subset I_{\mathcal{V}}$ for each target i [108]. A *relative distance function* is then defined as

$$\delta_j(\mathrm{y}_i) = \begin{cases} \frac{\|\mathrm{x}_j - \mathrm{y}_i\|}{\sum_{k \in \mathcal{B}(\mathrm{y}_i)} \|\mathrm{x}_k - \mathrm{y}_i\|}, & j \in \mathcal{B}(\mathrm{y}_i), \\ 1, & j \notin \mathcal{B}(\mathrm{y}_i), \end{cases} \qquad \forall j \in I_{\mathcal{V}}.$$

The assignment function is subsequently chosen as $a_{ij} = q(\delta_j(\mathrm{y}_i))$, for any $i \in I_{\mathcal{T}}$ and $j \in I_{\mathcal{V}}$, with

$$q(\delta) = \begin{cases} 1, & \delta \leq \Delta, \\ \frac{1}{1-2\Delta}[(1-\Delta) - \delta], & \Delta \leq \delta \leq 1 - \Delta, \\ 0, & 1 - \Delta \leq \delta, \end{cases}$$

where $\Delta \in [0, 1/2)$ is a prespecified parameter which can represent the capture radius in [108]. The second assignment scheme is more general than the first one.

Both of the assignment schemes described above suffer from two deficiencies: i) In the assigning procedure, they do not consider all the vehicles and targets at the same time. As a consequence, in the Voronoi-based assignment some targets may be remained unassigned to any vehicle, and in the proximity-based assignment some vehicles may be assigned to no target; ii) since the assignments are, to some extent, designed explicitly and are set to have a special structure, they may not constitute an optimal solution.

Before introducing the (implicit) *optimal assignment*, it is required to investigate the structure of task allocation. First, in order for every vehicle to be fully devoted to the tasks, the sum of its task assignments must be equal to one, i.e.

$$\sum_{i \in I_{\mathcal{T}}} a_{ij}(\mathrm{x}, \mathrm{y}) = 1, \quad \forall j \in I_{\mathcal{V}}, \tag{2.9}$$

where $\mathrm{x} = [\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_M]$ and $\mathrm{y} = [\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_N]$. As for the targets, there are two possibilities: i) $M \geq N$ and ii) $M \leq N$. In the first case, in order to increase the chances of task accomplishments, it is reasonable to act generously and over-assign the targets to the vehicles, as there is at least one vehicle for each target, i.e.

$$\sum_{j \in I_{\mathcal{V}}} a_{ij}(\mathrm{x}, \mathrm{y}) \geq 1, \quad \forall i \in I_{\mathcal{T}}. \tag{2.10}$$

when $M < N$, on the other hand, since the number of vehicles is less than the number of targets, in order to manage the resources efficiently and accomplish the tasks as much as possible, it is more preferable to act cautiously and under-assign the targets to the vehicles, i.e.

$$\sum_{j \in I_{\mathcal{V}}} a_{ij}(\mathrm{x}, \mathrm{y}) \leq 1, \quad \forall i \in I_{\mathcal{T}}. \tag{2.11}$$

Note that the equality in (2.10) holds when $M = N$.

Relations (2.9), (2.10) and (2.11) form a set of constraints that the optimal assignment $\{a_{ij}\}_{i \in I_{\mathcal{T}}, j \in I_{\mathcal{V}}}$ should satisfy. Denote by $\mathcal{A}_{I_{\mathcal{T}}, I_{\mathcal{V}}}$ the set of the assignments which satisfy these constraints, i.e.

$$\mathcal{A}_{I_{\mathcal{T}},I_{\mathcal{V}}} = \{A = (a_{ij}(\mathrm{x},\mathrm{y}))_{|I_{\mathcal{T}}|\times|I_{\mathcal{V}}|} : \mathcal{M}^{|I_{\mathcal{V}}|} \times \mathcal{M}^{|I_{\mathcal{T}}|} \rightarrow [0,1]^{|I_{\mathcal{T}}|\times|I_{\mathcal{V}}|} \; ; \; A^{\mathsf{T}}\mathbf{1}_{|I_{\mathcal{T}}|} = \mathbf{1}_{|I_{\mathcal{V}}|},$$
$$|I_{\mathcal{V}}| \geq |I_{\mathcal{T}}| \Rightarrow A\mathbf{1}_{|I_{\mathcal{V}}|} \geq \mathbf{1}_{|I_{\mathcal{T}}|}, \; |I_{\mathcal{V}}| \leq |I_{\mathcal{T}}| \Rightarrow A\mathbf{1}_{|I_{\mathcal{V}}|} \leq \mathbf{1}_{|I_{\mathcal{T}}|}\}, \tag{2.12}$$

where $\mathbf{1}_n$ represents an n dimensional column vector of ones.

It is straightforward ro show that a Voronoi-based assignment satisfies (2.9), while a proximity-based assignment satisfies (2.10) and (2.11).

2.2.2 Cooperative Receding Horizon Trajectory Construction

The cooperative receding horizon controller (CRHC) iteratively generates a set of headings, step sizes and optimal assignments for each vehicle such that the resulting trajectories guide the team toward maximizing the collected rewards. Let the time instants at which the CRHC is applied be denoted by $\{t_k\}_{k=0}^{\infty} \in [0, T]$. At any time instant t_k, an optimization problem is solved, which provides an estimation of the collectible rewards in the future. The problem composition is based on the current positions of the vehicles and targets, and also predicted future positions of the targets. The solution of the problem provides the optimal control input $\mathrm{u}^k = [\mathrm{u}_1(t_k), \mathrm{u}_2(t_k), \ldots, \mathrm{u}_M(t_k)]$ as well as the optimal assignment $\{a_{ij}(\mathrm{x}(t_{k+1}), \hat{\mathrm{y}}(t_{k+1}))\}_{i\in I_{\mathcal{T}}, j\in I_{\mathcal{V}}}$.

Let H_k be the CRHC planning horizon. Here, for the case of simplicity, take action horizon the same as planning horizon. Therefore $t_{k+1} = t_k + H_k$. Assuming that the control input $\mathrm{u}_j(t_k)$ is applied to vehicle j, for any $j \in I_{\mathcal{V}}$. Then the planned position of vehicle j at time t_{k+1} is given by

$$\mathrm{x}_j(t_{k+1}) = \mathrm{x}_j(t_k) + \mathrm{u}_j(t_k)H_k, \quad j \in I_{\mathcal{V}}.$$

Due to the current positions of targets and vehicles, and also the control input u^k, the predicted earliest possible time that vehicle j can visit target i is

$$\tau_{ij}(\mathrm{u}^k, t_k) = (t_k + H_k) + \frac{\|\mathrm{x}_j(t_k + H_k) - \hat{\mathrm{y}}_i(t_k + H_k)\|}{V_j + v}.$$

The above prediction will be true if the estimate $\hat{\mathrm{y}}_i(t_k + H_k)$ is exact, and vehicle j and target i move toward each other with maximum speed. Thus, if $a_{ij}(\mathrm{x}(t_{k+1}), \hat{\mathrm{y}}(t_{k+1}))$ is the optimal assignment, it is expected to remain unchanged until vehicle j visits target i. Therefore, one have

$$a_{ij}(\mathrm{x}(\tau_{ij}(\mathrm{u}^k, t_k)), \hat{\mathrm{y}}(\tau_{ij}(\mathrm{u}^k, t_k))) = a_{ij}(\mathrm{x}(t_{k+1}), \hat{\mathrm{y}}(t_{k+1})). \tag{2.13}$$

Accordingly, at the time t_{k+1} one can estimate the maximum reward which the team is expected to collect by the time the mission is accomplished. Denote this predicted expected reward by $\mathfrak{R}^{k+1}$.

In order to formulate $\mathfrak{R}^{k+1}$, let $\tilde{\rho}_{ij}(\mathrm{u}^k, t_k) = \rho_i[\tau_{ij}(\mathrm{u}^k, t_k)]$ and

$$\tilde{a}_{ij}(\mathrm{u}^k, t_k) = a_{ij}(\mathrm{x}(\tau_{ij}(\mathrm{u}^k, t_k)), \hat{\mathrm{y}}(\tau_{ij}(\mathrm{u}^k, t_k))).$$

From the definition of $\mathfrak{R}^{k+1}$, we have

$$\mathfrak{R}^{k+1}(\mathrm{u}^k, t_k) = \sum_{i \in I_{\mathcal{T}}(t_k)} \sum_{j \in I_{\mathcal{V}}(t_k)} R_i \tilde{\rho}_{ij}(\mathrm{u}^k, t_k) \tilde{a}_{ij}(\mathrm{u}^k, t_k), \tag{2.14}$$

where the time-dependency of the targets set and vehicles set is explicitly shown by using argument t_k in the corresponding index set. Note that $\mathfrak{R}^k(\mathrm{u}^k, t_k)$ is in fact an estimation performed at the current time, t_k, for the total reward that the team can expect at the next time instant t_{k+1} to be capable of collecting by the final time T.

Now, one can present the optimization problem P^k, as follows:

$$\begin{aligned} \max \quad & \mathfrak{R}^{k+1}(\mathrm{u}^k, t_k) \\ \text{s.t.} \quad & \tilde{A}(\mathrm{u}^k, t_k) \in \mathcal{A}^k, \\ & \mathrm{u}^k \in \mathcal{U}^k. \end{aligned} \tag{2.15}$$

where $\mathcal{A}^k = \mathcal{A}_{I_{\mathcal{T}}(t_k), I_{\mathcal{V}}(t_k)}$ and $\mathcal{U}^k = \{\mathrm{u} = [\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_M]$
$\in \mathbb{R}^{Md}; \mathrm{u}_j \in \mathbb{R}^d, \|\mathrm{u}_j\| \leq V_j, \forall j \in I_{\mathcal{V}}\}$ is the set of *admissible heading control.*

For convenience of notation, $\mathrm{x}_j(t_k)$, $\mathrm{y}_i(t_k)$ and $\hat{\mathrm{y}}_i(t_k)$ will hereafter be denoted by x_j^k, y_i^k and $\hat{\mathrm{y}}_i^k$, respectively, for any $i \in I_{\mathcal{T}}$ and $j \in I_{\mathcal{V}}$. Accordingly, the corresponding vectors are represented by x^k, y^k and $\hat{\mathrm{y}}^k$.

2.3 Stationary Analysis of Paths in Cooperative Receding Horizon

In this section, a theoretical analysis on the vehicles' trajectories in the presence of infinite number of temporally-rare targets appearing sequentially in the mission space is presented.

Assume the mission space $\mathcal{M}$ is compact, and let $\{T_i\}_{i=0}^{\infty}$ be a sequence of strictly increasing non-negative real numbers with $T_0 = 0$, where T_i represents the arrival time of the i^{th} target, for any $i \in \mathbb{N}$. Given a non-negative scalar Δ and and an integer $k \in \mathbb{N}$, the sequence $\{T_i\}_{i=0}^{\infty}$ is (Δ, k)-*rare*, if $T_{n+k} - T_n > \Delta$, for all $n \in \mathbb{N} \cup \{0\}$. As will be demonstrated in the sequel, for any $(\Delta, 1)$-rare sequence $\{T_i\}_{i=0}^{\infty}$, there exists a positive scalar Δ_0 such that if $\Delta > \Delta_0$, then with two vehicles in the mission space and at most two targets at the initial time T_0, the number of unaccomplished tasks will always remain less than or equal to two.

Definition 2. *The trajectory* $\mathrm{x}(t) = [\mathrm{x}_1(t), \mathrm{x}_2(t), \ldots, \mathrm{x}_M(t)] \in \mathcal{M}^d$ *is called a stationary trajectory if* $\|\mathrm{x}_j(t) - \mathrm{y}_i(t)\| \leq s_i$ *for some* $t_h \in [0, T]$*, refered to as the hitting time, and some indices* $i \in I_{\mathcal{T}}, j \in I_{\mathcal{V}}$.

Lemma 1. *Consider the vectors* $\mathrm{p}, \mathrm{q}, \mathrm{v} \in \mathbb{R}^d$ *and the set of non-negative real numbers*

$V, H, B \in \mathbb{R}_{>0}$. Assume that $\|\mathrm{v}\| < V$, and that $\alpha : [0, H] \to \mathbb{R}^d$ is a bounded vector-valued function defined over the interval $[0, H]$ with $\max_{t\in[0,H]} \|\alpha(t)\| \le B$. Define the set

$$\Omega_{\mathrm{q},H} = \{(\mathrm{w}, t) \in \mathbb{R}^d \times \mathbb{R}; t \in [0, H], \|\mathrm{w} - \mathrm{q}\| \le Vt\}, \tag{2.16}$$

which is a convex compact subset of $\mathbb{R}^d$, and let $\mathrm{z} : [0, H] \to \mathbb{R}^d$ be given by $\mathrm{z}(t) = \mathrm{p} + \mathrm{v}t + \frac{1}{2}\alpha(t)t^2$. Define also

$$H < \min\{\frac{\|\mathrm{p} - \mathrm{q}\|}{V + \|\mathrm{v}\|}, \frac{V - \|\mathrm{v}\|}{B}\}, \tag{2.17}$$

and

$$(\mathrm{w}^*, t^*) = \underset{(\mathrm{w},t)\in\Omega_{\mathrm{q},H}}{\operatorname{argmin}} \frac{1}{2}\|\mathrm{w} - \mathrm{p} - \mathrm{v}t\|^2. \tag{2.18}$$

Then

i*) $t^* = H$, $\|\mathrm{w}^* - \mathrm{q}\| = VH$;*

ii*)*

$$\|\mathrm{p} - \mathrm{q}\| - \|\mathrm{w}^* - \mathrm{z}(H)\| \ge f(H) \tag{2.19}$$

where $f(h) = h(V - \|v\| - \frac{1}{2}Bh)$, and

iii*) $\|\tilde{\mathrm{w}}^* - \mathrm{q}\| = VH$ and $\|\mathrm{p} - \mathrm{q}\| - \|\tilde{\mathrm{w}}^* - \mathrm{z}(H)\| \ge f(H)$, where*

$$\tilde{\mathrm{w}}^* = \operatorname{argmin}_{\|\mathrm{w}-\mathrm{q}\|\le VH} \|\mathrm{w} - \mathrm{p} - \mathrm{v}H\|. \tag{2.20}$$

Proof. Proof of part (i) Since (2.18) presents a convex optimization problem, it admits a unique solution which can be calculated using Karush-Kuhn-Tucker (KKT) theorem [128], [129]. Rewriting the problem in a standard form yields

$$\begin{aligned}
(\mathrm{w}^*, t^*) \quad &= \quad \operatorname{argmin} \tfrac{1}{2}\|\mathrm{w} - \mathrm{p} - \mathrm{v}t\|^2, \\
\text{s.t.} \quad & \|\mathrm{w} - \mathrm{q}\|^2 - (Vt)^2 \leq 0, \\
& t - H \leq 0, \\
& -t \leq 0.
\end{aligned}$$

Now, there exist non-negative real Lagrange multipliers μ_1, μ_2, and β such that

$$0 = \mathrm{w}^* - \mathrm{p} - \mathrm{v}t^* + \beta(\mathrm{w}^* - \mathrm{q}), \tag{2.21a}$$

$$0 = (\mathrm{p} + \mathrm{v}t^* - \mathrm{w}^*)^{\mathsf{T}}\mathrm{v} - \beta V^2 t^* + \mu_1 - \mu_2, \tag{2.21b}$$

$$0 = \beta(\|\mathrm{p}^* - \mathrm{q}\|^2 - (Vt^*)^2), \tag{2.21c}$$

$$0 = \mu_1(t^* - H), \tag{2.21d}$$

$$0 = -\mu_2 t^*. \tag{2.21e}$$

From the inequality (2.17) and on noting $\|\mathrm{v}\| < V$, it is concluded that

$$\begin{aligned} \|\mathrm{w}^* - \mathrm{p} - \mathrm{v}t^*\| &= \|(\mathrm{w}^* - \mathrm{q}) + (\mathrm{q} - \mathrm{p}) - \mathrm{v}t^*\| \\ &\geq \|\mathrm{q} - \mathrm{p}\| - \|\mathrm{w}^* - \mathrm{q}\| - \|\mathrm{v}\|t^* \\ &\geq \|\mathrm{q} - \mathrm{p}\| - Vt^* - \|\mathrm{v}\|t^* \\ &\geq \|\mathrm{q} - \mathrm{p}\| - (V + \|\mathrm{v}\|)H \\ &> 0. \end{aligned}$$

This means that $\mathrm{w}^* - \mathrm{p} - \mathrm{v}t^* \neq 0$, which implies $\beta > 0$ and $\mathrm{w}^* - \mathrm{q} \neq 0$. On the other hand, it results from the equations in (2.21), that $\|\mathrm{w}^* - \mathrm{q}\| = Vt^*$. If $t^* = 0$, then $\mathrm{w}^* - \mathrm{q} = 0$, which is a contradiction. Therefore $t^* \neq 0$, and consequently, $\mu_2 = 0$. From the equations in (2.21), one can also deduce

$$\begin{aligned} \mu_1 &= -(\mathrm{p} + \mathrm{v}t^* - \mathrm{w}^*)^{\mathsf{T}}\mathrm{v} + \beta V^2 t^* \\ &= -\beta(\mathrm{w}^* - \mathrm{q})^{\mathsf{T}}\mathrm{v} + \beta V^2 t^* \\ &= \beta(V^2 t^* - (\mathrm{w}^* - \mathrm{q})^{\mathsf{T}}\mathrm{v}). \end{aligned}$$

Now, using the Cauchy-Schwartz inequality [130], one arrives at

$$\begin{aligned} \mu_1 &\geq 2\beta(V^2 t^* - \|\mathrm{w}^* - \mathrm{q}\|\|\mathrm{v}\|) \\ &\geq 2\beta(V^2 t^* - V\|\mathrm{v}\|t^*) \\ &\geq 2\beta V t^*(V - \|\mathrm{v}\|) > 0. \end{aligned}$$

It follows from (2.21d) and the above inequality that $t^* = H$, and therefore $\|\mathrm{w}^* - \mathrm{q}\| = VH$. These completes the proof of part (i).

Proof of part (ii) Since

$$\mathrm{w}^* = \frac{\beta}{\beta+1}\mathrm{q} + \frac{1}{\beta+1}(\mathrm{p} + \mathrm{v}H),$$

and $\beta > 0$, thus w^* is on the line connecting the points q and $\mathrm{p}+\mathrm{v}H$ in $\mathbb{R}^d$, which yields in

$$\|(\mathrm{q} - \mathrm{w}^*) + (\mathrm{w}^* - \mathrm{p} - \mathrm{v}H)\| = \|\mathrm{q} - \mathrm{w}^*\| + \|\mathrm{w}^* - \mathrm{p} - \mathrm{v}H\|.$$

This results in

$$\begin{aligned}
\|\mathrm{q} - \mathrm{p}\| &= \|(\mathrm{q} - \mathrm{p} - \mathrm{v}H) + \mathrm{v}H\| \\
&\geq \|\mathrm{q} - \mathrm{p} - \mathrm{v}H\| - \|\mathrm{v}H\| \\
&= \|(\mathrm{q} - \mathrm{w}^*) + (\mathrm{w}^* - \mathrm{p} - \mathrm{v}H)\| - H\|\mathrm{v}\| \\
&= \|\mathrm{q} - \mathrm{w}^*\| + \|\mathrm{w}^* - \mathrm{p} - \mathrm{v}H\| - \|\mathrm{v}\|H \\
&= VH + \|\mathrm{w}^* - \mathrm{p} - \mathrm{v}H\| - \|\mathrm{v}\|H,
\end{aligned} \tag{2.22}$$

and hence

$$\|\mathrm{q} - \mathrm{p}\| - \|\mathrm{w}^* - \mathrm{p} - \mathrm{v}H\| \geq (V - \|\mathrm{v}\|)H. \tag{2.23}$$

On the other hand, it results from the relations $\mathrm{z}(t) - \mathrm{p} - \mathrm{v}t = \frac{1}{2}\alpha(t)t^2$ and $\max_{t\in[0,H]} \|\alpha(t)\| \leq$

B, that $\|\mathrm{z}(H) - \mathrm{p} - \mathrm{v}H\| \le \frac{1}{2}BH^2$, which along with the relation

$$\|\mathrm{w}^* - \mathrm{z}(H)\| \le \|\mathrm{w}^* - \mathrm{p} - \mathrm{v}H\| + \|\mathrm{p} + \mathrm{v}H - \mathrm{z}(H)\|,$$

leads to

$$\|\mathrm{w}^* - \mathrm{z}(H)\| \le \|\mathrm{w}^* - \mathrm{p} - \mathrm{v}H\| + \tfrac{1}{2}BH^2. \tag{2.24}$$

By combining (2.23) and (2.24), one arrives at

$$\|\mathrm{p} - \mathrm{q}\| - \|\mathrm{w}^* - \mathrm{z}(H)\| \ge H(V - \|\mathrm{v}\| - \frac{1}{2}BH). \tag{2.25}$$

This concludes the proof of part (ii).

Proof of part (iii) As the first step of the proof, note that

$$\tilde{\mathrm{w}}^* = \mathrm{argmin}_{\|\mathrm{w}-\mathrm{q}\| \le VH} \ \frac{1}{2}\|\mathrm{w} - \mathrm{p} - \mathrm{v}H\|^2. \tag{2.26}$$

Now, let

$$\mathrm{w}^*(t) = \underset{\|\mathrm{w}-\mathrm{q}\| \le Vt}{\mathrm{argmin}} \ \frac{1}{2}\|\mathrm{w} - \mathrm{p} - \mathrm{v}t\|^2.$$

for some $t \in [0, H]$, which means $\tilde{\mathrm{w}}^* = \mathrm{w}^*(H)$. Moreover,

$$\min_{(\mathrm{w},t) \in \Omega_{\mathrm{q},H}} \frac{1}{2}\|\mathrm{w} - \mathrm{p} - \mathrm{v}t\|^2 = \min_{0 \le t \le H} \min_{\|\mathrm{w}-\mathrm{q}\| \le Vt} \ \frac{1}{2}\|\mathrm{w} - \mathrm{p} - \mathrm{v}t\|^2.$$

Thus

$$\frac{1}{2}\|\mathrm{w}^* - \mathrm{p} - \mathrm{v}t^*\|^2 \leq \frac{1}{2}\|\tilde{\mathrm{w}}^* - \mathrm{p} - \mathrm{v}H\|^2. \tag{2.27}$$

From (2.26) and on noting that $t^* = H$ and $\|\mathrm{w}^* - q\| = VH$ (according to part (i) of the lemma), it is concluded that

$$\frac{1}{2}\|\mathrm{w}^* - \mathrm{p} - \mathrm{v}t^*\|^2 \geq \frac{1}{2}\|\tilde{\mathrm{w}}^* - \mathrm{p} - \mathrm{v}H\|^2. \tag{2.28}$$

It follows from (2.27) and (2.28) that

$$\frac{1}{2}\|\mathrm{w}^* - \mathrm{p} - \mathrm{v}t^*\|^2 = \frac{1}{2}\|\tilde{\mathrm{w}}^* - \mathrm{p} - \mathrm{v}H\|^2. \tag{2.29}$$

Since (2.18) and (2.20) are strictly convex, thus $\mathrm{w}^* = \tilde{\mathrm{w}}^*$. The proof of part (iii) follows immediately from parts (i) and (ii). □

Remark 2. *The function $f(h) = h(V - \|\mathrm{v}\| - \frac{1}{2}Bh)$ introduced in Lemma 1 is a concave quadratic function which is: (*i*) non-negative only in the interval $I = [0, 2(V - \|\mathrm{v}\|)/B]$; (*ii*) zero only at the endpoints of interval; (*iii*) strictly increasing in interval $[0, (V - \|\mathrm{v}\|)/B]$, and (*iv*) attains its maximum at the midpoint of the interval I. Therefore, $f(H)$ is positive, and if*

$$\|\mathrm{p} - \mathrm{p}\| \geq \frac{V^2 - \|\mathrm{v}\|^2}{B},$$

then the function f takes its maximum value at H, i.e.

$$f(h) \leq f(H) = \frac{(V - \|\mathrm{v}\|)^2}{2B}, \quad \forall h \in \mathbb{R}.$$

For simplicity of the analysis, it is assumed hereafter that $f_1 = f_2 = 1$ and $V_1 = V_2 = V$.

Theorem 1 (Convergence of the scheme)**.** *Consider the optimal cooperative receding horizon problem presented in (2.15), and let $(M, N) \in \{(2,1), (2,2)\}$. Let also*

$$H_k = \min\{\min_{j \in I_{\mathcal{V}}} \frac{||x_j(t_k) - y_i(t_k)|| - \frac{1}{2}s_i}{v + V}, \frac{V - v}{3B}\}. \tag{2.30}$$

and assume that $v < V$. Then for any initial choice of A satisfying (2.9), (2.10) and (2.11), the cooperative receding horizon algorithm is finite-time convergent, i.e., the vehicles reach the targets in finite time.

Proof. During the time interval $[t_k, t_k + H_k]$ the control input u^k is constant. Therefore, the assignment maps $\tilde{a}_{ij}(\mathrm{u}^k, t_k)$, $i \in I_{\mathcal{T}}, j \in I_{\mathcal{V}}$, are constant in this time interval. Denote these assignment maps by a_{ij}^{k+1}, and consider the two possible scenarios $(M, N) = (2, 1)$ and $(M, N) = (2, 2)$ separately.

Case I: In this case, $I_{\mathcal{T}} = \{1\}$ and $I_{\mathcal{V}} = \{1, 2\}$. It results from equations (2.9), (2.10) and (2.11), that $a_{11}^{k+1} = a_{21}^{k+1} = 1$. Hence

$$\sum_{i \in I_{\mathcal{T}}} \sum_{j \in I_{\mathcal{V}}} R_i \tilde{\rho}_{ij}(\mathrm{u}^k, t_k) \tilde{a}_{ij}(\mathrm{u}^k, t_k) = R_1 \rho_{11}(\mathrm{u}^k, t_k) + R_1 \rho_{12}(\mathrm{u}^k, t_k).$$

(Recall that R_1 is the maximum reward for the target). Since

$$\rho_{ij}(\mathrm{u}^k, t_k) = 1 - \frac{1}{T}(t_k + H_k + \frac{\|\mathrm{x}_j(t_k + H_k) - \hat{\mathrm{y}}_i(t_k + H_k)\|}{V + v}). \tag{2.31}$$

One can write

$$\sum_{i \in I_{\mathcal{T}}} \sum_{j \in I_{\mathcal{V}}} R_i \tilde{\rho}_{ij}(\mathrm{u}^k, t_k) \tilde{a}_{ij}(\mathrm{u}^k, t_k) = 2R_1(1 - \frac{1}{T}(t_k + H_k)) - \frac{R_1}{T(V + v)} \sum_{j \in I_{\mathcal{V}}} \|\mathrm{x}_j(t_k + H_k) - \hat{\mathrm{y}}(t_k + H_k)\|.$$

Since $2R_1(1 - \frac{1}{T}(t_k + H_k))$ is constant, the optimization problem (2.15) can be simplified to

$$\begin{aligned} \min \quad & \textstyle\sum_{j \in I_{\mathcal{V}}} \|\mathrm{x}_j^{k+1} - \hat{\mathrm{y}}^{k+1}\| \\ \text{s.t.} \quad & \mathrm{x}_j^{k+1} = \mathrm{x}_j^k + \mathrm{u}_j^k, \quad j \in I_{\mathcal{V}}, \\ & \mathrm{u}^k \in \mathcal{U}^k. \end{aligned} \tag{2.32}$$

From the definition of $\mathcal{U}^k$, the above problem can be reformulated as

$$\begin{aligned} \min \quad & \|\mathrm{w}_1^k - \hat{\mathrm{y}}^{k+1}\| + \|\mathrm{w}_2^k - \hat{\mathrm{y}}^{k+1}\| \\ \text{s.t.} \quad & \|\mathrm{w}_j^k - \mathrm{x}_j^k\| \leq V H_k, \quad j \in I_{\mathcal{V}}, \end{aligned}$$

where $\mathrm{w}_j^k := \mathrm{x}_j^k + \mathrm{u}_j^k$, $j = 1, 2$. The above problem is equivalent to

$$\mathrm{w}_j^{*,k} = \mathrm{argmin}_{\|\mathrm{w}_j^k - \mathrm{x}_j^k\| \leq V H_k} \|\mathrm{w}_j^k - \hat{\mathrm{y}}^{k+1}\|, \quad j \in I_{\mathcal{V}}.$$

Using Lemma 1, equation (2.4) and Assumption 2, the following relation is obtained

$$\|\mathrm{x}_j^k - \mathrm{y}^k\| - \|\mathrm{x}_j^{k+1} - \mathrm{y}^{k+1}\| \geq f(H_k), \quad j \in I_{\mathcal{V}}.$$

Define $J^k = \|\mathrm{x}_1^k - \mathrm{y}^k\| + \|\mathrm{x}_2^k - \mathrm{y}^k\|$. Then

$$J^k - J^{k+1} \geq 2f(H_k). \tag{2.33}$$

If the trajectory $[\mathrm{x}_1, \mathrm{x}_2] \in \mathbb{R}^{2d}$ is non-stationary, i.e., for all $k \in \mathbb{N}$ and $j \in I_{\mathcal{V}}$, $\|\mathrm{x}_j^k - \mathrm{y}^k\| \geq s$, then

$$H_k \geq s = \min \left\{ \min_{i \in I_{\mathcal{T}}} \frac{s_i}{2(v + V)}, \frac{V - v}{3B} \right\} > 0. \tag{2.34}$$

Thus, it results from Remark 2, that $J^k - J^{k+1} \geq 2f(s) > 0$. It can be prove by induction that

$$J^0 - J^m \geq 2mf(s), \quad m \in \mathcal{N}. \tag{2.35}$$

Since J^0 and $f(s)$ are strictly positive, hence $\lim_{m\to\infty} J^m = -\infty$ which contradicts the fact that $J^m \geq 0$, for all $m \in \mathbb{N}$. Thus, the trajectory is stationary, i.e., there exist a finite k and some $j \in I_{\mathcal{V}}$ such that $\|\mathrm{x}_j^k - \mathrm{y}^k\| < s$. This completes the proof.

Case II: In this case, $I_{\mathcal{T}} = \{1, 2\}$ and $I_{\mathcal{V}} = \{1, 2\}$. From (2.12), it is straightforward to show that in this special case $a_{12}^{k+1} = a_{21}^{k+1} = 1 - a_{11}^{k+1}$ and $a_{22}^{k+1} = a_{11}^{k+1}$ or equivalently

$$\mathcal{A}^k = \Big\{ \begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix}; a \in [0, 1] \Big\}.$$

As a result

$$\sum_{i\in I_{\mathcal{T}}} \sum_{j\in I_{\mathcal{V}}} R_i \tilde{\rho}_{ij}(\mathrm{u}^k, t_k)\tilde{a}_{ij}(\mathrm{u}^k, t_k) = R_1\rho_{11}(\mathrm{u}^k, t_k)a_{11}^{k+1} + R_1\rho_{12}(\mathrm{u}^k, t_k)(1 - a_{11}^{k+1}) + \\ R_2\rho_{21}(\mathrm{u}^k, t_k)(1 - a_{11}^{k+1}) + R_2\rho_{22}(\mathrm{u}^k, t_k)a_{11}^{k+1}.$$

It follows from the above result and equation (2.31) that

$$\sum_{i\in I_{\mathcal{T}}} \sum_{j\in I_{\mathcal{V}}} R_i \tilde{\rho}_{ij}(\mathrm{u}^k, t_k)\tilde{a}_{ij}(\mathrm{u}^k, t_k) = (R_1 + R_2)(1 - \frac{1}{T}(t_k + H_k)) \\ - \frac{1}{T(V + v)} J(\mathrm{x}_1^{k+1}, \mathrm{x}_2^{k+1}, \hat{\mathrm{y}}_1^{k+1}, \hat{\mathrm{y}}_2^{k+1}, a^{k+1}),$$

where the function J is defined below

$$
\begin{aligned}
J(\mathrm{x}_1, \mathrm{x}_2, \mathrm{y}_1, \mathrm{y}_2, a) = {} & R_1 a\|\mathrm{x}_1 - \mathrm{y}_1\| + R_1(1-a)\|\mathrm{x}_2 - \mathrm{y}_1\| \\
& + R_2(1-a)\|\mathrm{x}_1 - \mathrm{y}_2\| + R_2 a\|\mathrm{x}_2 - \mathrm{y}_2\|.
\end{aligned} \tag{2.36}
$$

Since $(R_1 + R_2)(1 - \frac{1}{T}(t_k + H_k))$ is constant, the optimization problem (2.15) can be written as

$$
\begin{aligned}
\min \quad & J(\mathrm{x}_1^{k+1}, \mathrm{x}_2^{k+1}, \hat{\mathrm{y}}_1^{k+1}, \hat{\mathrm{y}}_2^{k+1}, a^{k+1}) \\
\text{s.t.} \quad & \mathrm{x}_j^{k+1} = \mathrm{x}_j^k + H_k \mathrm{u}_j^k, \quad j \in I_{\mathcal{V}}, \\
& a^{k+1} \in [0,1], \\
& \mathrm{u}^k \in \mathcal{U}^k.
\end{aligned} \tag{2.37}
$$

Similarly to the previous case, the above problem can be reformulated as

$$
\begin{aligned}
\min \quad & J(\mathrm{w}_1^k, \mathrm{w}_2^k, \hat{\mathrm{y}}_1^{k+1}, \hat{\mathrm{y}}_2^{k+1}, a) \\
\text{s.t.} \quad & \|\mathrm{w}_j^k - \mathrm{x}_j^k\| \leq V H_k, \quad j \in I_{\mathcal{V}}, \\
& a^{k+1} \in [0,1].
\end{aligned} \tag{2.38}
$$

Let $(\mathrm{w}_1^{*,k}, \mathrm{w}_2^{*,k}, a^{*,k+1})$ be the solution of (2.38). Also, for any $a \in [0,1]$, let

$$
\begin{aligned}
(\mathrm{w}_{1,a}^*, \mathrm{w}_{2,a}^*) = \quad & \operatorname{argmin} \quad J(\mathrm{w}_1, \mathrm{w}_2, \hat{\mathrm{y}}_1^{k+1}, \hat{\mathrm{y}}_2^{k+1}, a) \\
& \text{s.t.} \quad \|\mathrm{w}_j - \mathrm{x}_j^k\| \leq V H_k, \quad j = 1, 2.
\end{aligned} \tag{2.39}
$$

Note that for $a = 0$,

$$J(\mathrm{x}_1, \mathrm{x}_2, \mathrm{y}_1, \mathrm{y}_2, 0) = R_1\|\mathrm{x}_2 - \mathrm{y}_1\| + R_2\|\mathrm{x}_1 - \mathrm{y}_2\|, \tag{2.40}$$

and for $a = 1$,

$$J(\mathrm{x}_1, \mathrm{x}_2, \mathrm{y}_1, \mathrm{y}_2, 1) = R_1\|\mathrm{x}_1 - \mathrm{y}_1\| + R_2\|\mathrm{x}_2 - \mathrm{y}_2\|. \tag{2.41}$$

It results from equations (2.40) and (2.41) that for any $a \in \{0, 1\}$, the optimization problem (2.38)

$$\begin{aligned} (\mathrm{w}_{1,a}^{*,k}, \mathrm{w}_{2,a}^{*,k}) = \quad & \operatorname{argmin} \quad R_{2-a}\|\mathrm{w}_1^k - \hat{\mathrm{y}}_{2-a}^{k+1}\| + R_{a+1}\|\mathrm{w}_2^k - \hat{\mathrm{y}}_{a+1}^{k+1}\|, \\ & \text{s.t.} \quad \|\mathrm{w}_j^{k+1} - \mathrm{x}_j^k\| \le V H_k, \quad j \in I_{\mathcal{V}}. \end{aligned} \tag{2.42}$$

which can be decomposed to the following two optimization problems

$$\mathrm{w}_{1,a}^{*,k} = \operatorname{argmin}_{\|\mathrm{w}_1^k - \mathrm{x}_j^k\| \le V H_k} R_{2-a}\|\mathrm{w}_1^k - \hat{\mathrm{y}}_{2-a}^{k+1}\|,$$

and

$$\mathrm{w}_{2,a}^{*,k} = \operatorname{argmin}_{\|\mathrm{w}_2^k - \mathrm{x}_j^k\| \le V H_k} R_{1+a}\|\mathrm{w}_2^k - \hat{\mathrm{y}}_{1+a}^{k+1}\|.$$

Using Lemma 1, equation (2.4) and Assumption 2, it is concluded that

$$\|\mathrm{x}_1^k - \mathrm{y}_{2-a}^k\| - \|\mathrm{w}_{1,a}^* - \mathrm{y}_{2-a}^{k+1}\| \geq f(H_k), \quad j \in I_{\mathcal{V}},$$

and

$$\|\mathrm{x}_2^k - \mathrm{y}_{1+a}^k\| - \|\mathrm{w}_{2,a}^* - \mathrm{y}_{1+a}^{k+1}\| \geq f(H_k), \quad j \in I_{\mathcal{V}}.$$

Therefore, for any $a \in \{0, 1\}$, the following relation holds

$$J(\mathrm{x}_1^k, \mathrm{x}_2^k, \mathrm{y}_1^k, \mathrm{y}_2^k, a) - J(\mathrm{w}_{1,a}^*, \mathrm{w}_{2,a}^*, \mathrm{y}_1^{k+1}, \mathrm{y}_2^{k+1}, a) \geq (R_1 + R_2) f(H_k). \tag{2.43}$$

On the other hand, from the definition of J, for any $\mathrm{x}_1, \mathrm{x}_2, \mathrm{y}_1, \mathrm{y}_2 \in \mathbb{R}^d$ and $a \in [0, 1]$, one can write

$$J(\mathrm{x}_1, \mathrm{x}_2, \mathrm{y}_1, \mathrm{y}_2, a) = (1 - a) J(\mathrm{x}_1, \mathrm{x}_2, \mathrm{y}_1, \mathrm{y}_2, 0) + a J(\mathrm{x}_1, \mathrm{x}_2, \mathrm{y}_1, \mathrm{y}_2, 1). \tag{2.44}$$

Define $\bar{a}^k = \mathrm{argmin}_{a \in \{0,1\}} J(\mathrm{x}_1^k, \mathrm{x}_2^k, \mathrm{y}_1^k, \mathrm{y}_2^k, a)$ to obtain

$$J(\mathrm{x}_1^k, \mathrm{x}_2^k, \mathrm{y}_1^k, \mathrm{y}_2^k, a^k) \geq J(\mathrm{x}_1^k, \mathrm{x}_2^k, \mathrm{y}_1^k, \mathrm{y}_2^k, \bar{a}^k). \tag{2.45}$$

Since for any $i \in I_{\mathcal{T}}$ the relation $|\mathrm{y}_i^k - \hat{\mathrm{y}}_i^k| \leq \frac{1}{2} B H_k^2$ holds, using the definition of J

(equation (2.36)), it is straightforward to show that

$$|J(\mathrm{x}_1, \mathrm{x}_2, \hat{\mathrm{y}}_1^k, \hat{\mathrm{y}}_2^k, a) - J(\mathrm{x}_1, \mathrm{x}_2, \mathrm{y}_1^k, \mathrm{y}_2^k, a)| \leq \frac{1}{2}(R_1 + R_2)BH_k^2, \tag{2.46}$$

for all $\mathrm{x}_1, \mathrm{x}_2 \in \mathbb{R}^d$ and $a \in [0, 1]$. Therefore, if $J^k = J(\mathrm{x}_1^k, \mathrm{x}_2^k, \hat{\mathrm{y}}_1^k, \hat{\mathrm{y}}_2^k, a^k)$, then for every $k \geq 0$

$$J^k \geq J(\mathrm{x}_1^k, \mathrm{x}_2^k, \mathrm{y}_1^k, \mathrm{y}_2^k, a^k) - \frac{1}{2}(R_1 + R_2)BH_k^2, \tag{2.47}$$

and consequently

$$J^k \geq J(\mathrm{x}_1^k, \mathrm{x}_2^k, \mathrm{y}_1^k, \mathrm{y}_2^k, \bar{a}^k) - \frac{1}{2}(R_1 + R_2)BH_k^2. \tag{2.48}$$

Thus, it results from (2.43) that

$$J^k \geq J(\mathrm{w}_{1,\bar{a}^k}^*, \mathrm{w}_{2,\bar{a}^k}^*, \mathrm{y}_1^{k+1}, \mathrm{y}_2^{k+1}, \bar{a}^k) + (R_1 + R_2)(f(H_k) - \frac{1}{2}BH_k^2). \tag{2.49}$$

Now, (2.46) and (2.49) yield

$$J^k \geq J(\mathrm{w}_{1,\bar{a}^k}^*, \mathrm{w}_{2,\bar{a}^k}^*, \hat{\mathrm{y}}_1^{k+1}, \hat{\mathrm{y}}_2^{k+1}, \bar{a}^k) + (R_1 + R_2)(f(H_k) - BH_k^2). \tag{2.50}$$

Since $(\mathrm{w}_1^{*,k}, \mathrm{w}_2^{*,k}, a^{*,k+1})$ is the solution of (2.38), it can be concluded that

$$J(\mathrm{w}^*_{1,\bar{a}^k}, \mathrm{w}^*_{2,\bar{a}^k}, \hat{\mathrm{y}}_1^{k+1}, \hat{\mathrm{y}}_2^{k+1}, \bar{a}^k) \geq J^{k+1}, \tag{2.51}$$

and therefore

$$J^k \geq J^{k+1} + (R_1 + R_2)(f(H_k) - BH_k^2). \tag{2.52}$$

If the trajectory $[\mathrm{x}_1, \mathrm{x}_2] \in \mathbb{R}^{2d}$ is non-stationary (i.e., for $k \in \mathbb{N}$, $i \in I_{\mathcal{T}}$ and $j \in I_{\mathcal{V}}$, the relation $\|\mathrm{x}_j^k - \mathrm{y}_i^k\| \geq s_i$ holds), then

$$H_k \geq s = \min\left\{\min_{i \in I_{\mathcal{T}}} \frac{s_i}{2(v+V)}, \frac{V-v}{3B}\right\} > 0. \tag{2.53}$$

Let $g(h) = f(h) - Bh^2$ or equivalently $g(h) = h(V - v - \frac{3}{2}Bh)$. One can show that similar to f, the function g is also a concave quadratic function which is: (i) non-negative only in the interval $I = [0, 2(V-v)/(3B)]$; (ii) zero only at the endpoints of the interval I; (iii) strictly increasing in the interval $[0, (V-v)/(3B)]$, and (iv) attains its maximum at the midpoint of the interval I. Therefore, $g(H_k) \geq g(s) > 0$, which implies that $J^k - J^{k+1} \geq (R_1 + R_2)g(s) > 0$. It can be shown by induction that for all $m \in \mathcal{N}$

$$J^0 - J^m \geq m(R_1 + R_2)g(s), \quad m \in \mathcal{N}. \tag{2.54}$$

Thus, the inequalities $g(s) > 0, R_1 + R_2 > 0$ yield

$$\lim_{m\to\infty} J^m = -\infty, \tag{2.55}$$

which contradicts the fact that $J^m \geq 0$, for all $m \in \mathbb{N}$. This means that the trajectory is stationary, i.e., there is a finite k and some $i \in I_{\mathcal{T}}, j \in I_{\mathcal{V}}$, such that $\|\mathrm{x}_j^k - \mathrm{y}_i^k\| < s$. This completes the proof. □

Corollary 1. *Let t_h be the hitting time introduced in Definition 2, and*

$$H_{\min} := \min\left\{\min_{i\in I_{\mathcal{T}}} \frac{s_i}{2(v+V)}, \frac{V-v}{3B}\right\} > 0.$$

Then

$$t_h \leq \frac{2\ \mathrm{diam}(\mathcal{M})}{V-v}, \tag{2.56}$$

where $\mathrm{diam}(\mathcal{M}) = \sup_{\mathrm{m}_1,\mathrm{m}_2\in\mathcal{M}} \|\mathrm{m}_1 - \mathrm{m}_2\|$ *is the diameter of the set* $\mathcal{M}$.

Proof. The proof follows directly from equations (2.35) and (2.54), and the fact that $f(s) > s(V-v)/B$ and $g(s) > s(V-v)/(3B)$. □

It follows from Corollary 1 that if $\Delta > 2\mathrm{diam}(\mathcal{M})/(V-v)$, then at any time instant before the arrival of the next target, at least one of the present targets will be visited. Therefore, the number of unaccomplished tasks is non-increasing and remains less than or equal to two.

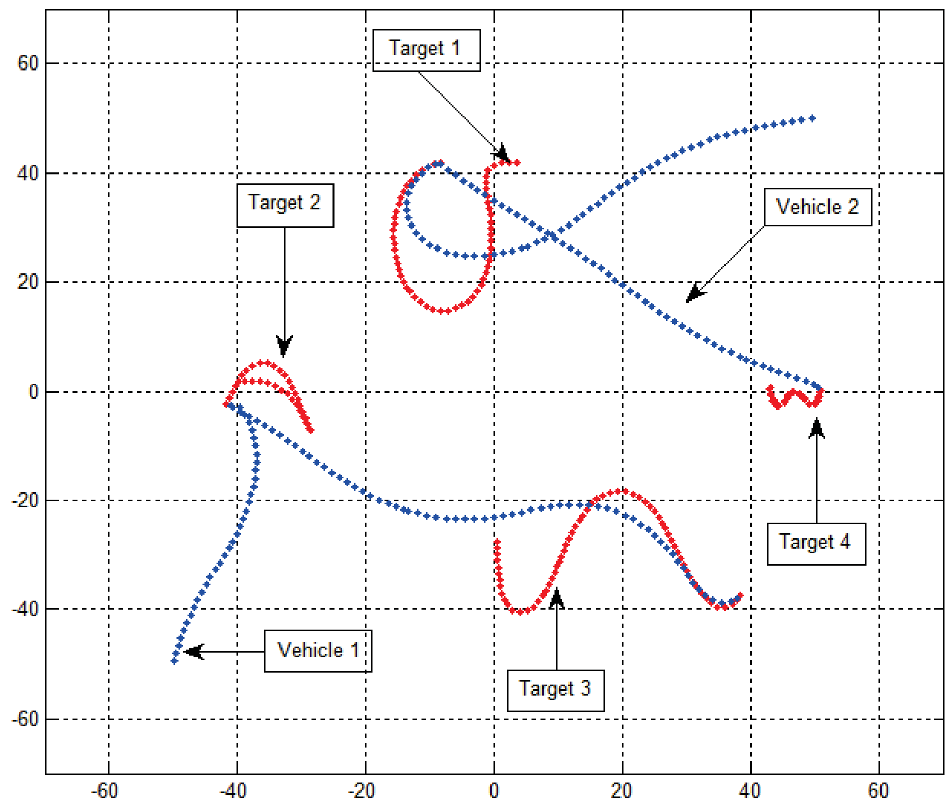

Figure 2.1: The target tracking for the vehicles and sequentially arriving targets of Example 1, using the proposed control strategy.

2.4 Simulation Results

In this section, simulations are performed for an example involving two vehicles and a set of targets arriving sequentially in the mission space illustrated in Fig. 2.1.

Example 1. *Let the mission space, fig 2.1 be a* $\mathcal{M} = [-60, 60] \times [-60, 60]$ *closed convex set in the 2D plane. Let also two vehicles be inside the mission space. Assume that initially there exist two targets in* $\mathcal{M}$ *with some* a priori *unknown trajectories satisfying Assumptions 1 and 2, and that new targets arrive sequentially in the mission space* *afterwards. For generality, the targets trajectories are chosen randomly. The maximum magnitude of the velocity vector for the vehicles is assumed to be* $V = 5$*, and the upper bound on the magnitude of targets velocities is given as* $v = 4$*, while the bound introduced in Assumption 2 is chosen as* $B = 1$*. The arrival time of targets, on the other hand, is given by the sequence* $\{T_i\}_{i=1}^{\infty} = \{0, 5.0851, 9.2216, \ldots\}$ *(which is generated randomly). To illustrates the results, a snapshot is shown at time* $T = 20$*, and four targets, including initial targets, arrive in the mission space by this time. At the time instants when there is no target in the mission space, the vehicles remain at their last position until the arrival of new targets. Using the cooperative control approach developed in this chapter, the results obtained in Fig. 2.1 are obtained. As it can be observed from this figure, all targets are visited by the vehicles and the tasks are accomplished accordingly. This demonstrates the efficacy of the proposed strategy for the system given in this example.*

Chapter 3

Stability Analysis of Dynamic Decision-Making for Vehicle Heading Control

In this chapter, the problem of dynamic decision making for vehicle heading control to intercept moving targets is investigated. It is assumed that the targets arrive in the mission space sequentially. More precisely, there exist infinite number of targets that arrive the mission space one by one. The arrival times and positions of the targets are modeled using stochastic models. Furthermore, targets are assumed to move with *a priori* unknown dynamics and *a priori* unknown trajectories. Due to the probabilistic nature of the problem, it is desired to use a model predictive approach to control the heading of the vehicle. A reward allocation strategy is adopted for dynamic decision making and control design in order to move the vehicle toward the targets. Finite-time convergence analysis is presented for the case where the arrivals of targets occur sufficiently infrequently.

3.1 Preliminaries and Notations

3.1.1 Notations

Throughout this chapter, $\mathbb{N}, \mathbb{Z}, \mathbb{R}$ denote the set of natural numbers, integers and real numbers, respectively, and the index inequalities in $\mathbb{R}_{\geq 0}, \mathbb{R}_{>0}, \mathbb{Z}_{\geq 0}$ and other sets represent inequalities imposed over the elements of the corresponding set. Also, $\mathbb{N}_n$ denotes natural numbers less than or equal to n. The symmetric difference of a pair of sets A and B is defined as $(A \cap B^c) \cup (A^c \cap B)$ and denoted by $A \Delta B$, where the superscript "c" represents the complement operator. Given a set A, the Kronecker delta function, denoted by δ, maps $A \times A$ to $\{0, 1\}$, where $\delta(a, b) = 1$ if and only if $a = b$. For simplicity of notation, $\delta(a, b)$ is δ_{ab}. The indicator function for any subset B of A, denoted by $\mathbf{1}_B$, is a function from A to $\{0, 1\}$, defined as

$$\mathbf{1}_B(x) := \begin{cases} 1, & \text{if } x \in B, \\ 0, & \text{if } x \notin B. \end{cases}$$

Let $\mathcal{I}$ be the index set. Then $(a_i)_{i\in\mathcal{I}}$ represents a point in $A^{\mathcal{I}}$ with entries a_i. If $\mathcal{J}$ is a non-empty index set such that $\mathcal{J} \subseteq \mathcal{I}$, then for any point in $\mathrm{a} \in A^{\mathcal{I}}$, $\mathrm{a}_{\bullet\mathcal{J}}$ represents a point in $A^{\mathcal{J}}$ which is obtained by eliminating the entries with indices not listed in $\mathcal{J}$.

The d dimensional Euclidean space is denoted by $\mathbb{R}^d$. Moreover, $\mathbf{0}$ and $\mathbf{1}$ represents all-zero and all-one vectors in $\mathbb{R}^d$. The notation $\mathrm{a} \geq 0$ says that all entries of a are non-negative. For any compact set $\mathcal{M} \subset \mathbb{R}^d$, the *diameter* of $\mathcal{M}$, denoted by $\mathrm{diam}(\mathcal{M})$, is defined as

$$\mathrm{diam}(\mathcal{M}) = \sup\{\|\mathrm{x} - \mathrm{y}\| \mid \mathrm{x}, \mathrm{y} \in \mathcal{M}\}. \tag{3.1}$$

For any scalar $r \in \mathbb{R}_{\geq 0}$ and any point $\mathrm{x} \in \mathbb{R}^d$, the closed ball with radius r centered at x is defined as

$$\mathscr{B}(\mathrm{x}, r) = \{\mathrm{y} \in \mathbb{R}^d \mid \|\mathrm{x} - \mathrm{y}\| \leq r\}. \tag{3.2}$$

For any vector $\mathrm{v} \in \mathbb{R}^d$, the perpendicular complement of v, denoted by $\mathrm{v}^{\perp}$, is a $(d-1)$-dimensional subspace of $\mathbb{R}^d$ defined as

$$\mathrm{v}^{\perp} := \{\mathrm{y} \in \mathbb{R}^d \mid \mathrm{v}^{\mathsf{T}}\mathrm{y} = 0\}. \tag{3.3}$$

3.1.2 Mathematical Preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where Ω, $\mathcal{F}$, $\mathbb{P}$ represents sample space, σ-algebra of events and probability measure, respectively. Then, the sequence of random variables (vectors) $\{X_n\}_{n\in\mathbb{N}}$ converges $\mathbb{P}$-almost surely to a random variable (vector) X if $\mathbb{P}(X_n \to_{n\to\infty} X) = 1$. This convergence is denoted by

$$X_n \to_{n\to\infty} X, \quad \mathbb{P}\text{-a.s.} \tag{3.4}$$

Definition 3. *A stochastic process* $\mathrm{N} = \{N(t)\}_{t\geq 0}$ *is a renewal process if there exists a sequence of independent identically distributed (i.i.d.) non-negative random variables* $\{X_m\}_{n\in\mathbb{N}}$ *such that* $N(t) = \max\{n \in \mathbb{Z}_{\geq 0} \mid \sum_{i=1}^{n} X_i \leq t\}$. *Without loss of generality, let* $\sum_{i=1}^{0} X_i := 0$.

Theorem 2. *[131] In a renewal process, let* $\mu = 1/\mathbb{E}X_i$, *where* $\mathbb{E}$ *denotes the expectation operator. Then*

$$\frac{N(t)}{t} \to \mu \quad \mathbb{P}\text{-}a.s. \tag{3.5}$$

3.2 Problem Formulation

Consider a closed convex subset of $\mathbb{R}^d$ as the *mission space*, denoted by $\mathcal{M}$, and a vehicle inside it with dynamics described by

$$\dot{\mathrm{x}}(t) = \mathrm{u}(t) = V(t)\mathrm{d}(t), \tag{3.6}$$

where $V(t) \in [0, V_{\max}]$ is the control input for the magnitude of the velocity vector, and $\mathrm{d}(t) \in \mathbb{S}^{d-1} = \{\mathrm{d} \in R^d; \|\mathrm{d}\| = 1\}$ is the control input for its direction.

Assume the mission starts at time $t = 0$ and let a sequence of moving targets arrive in the mission space randomly in time and space. The arrival process can be described using a spatial model and a temporal model. Let $N_0 \in \mathbb{N}_0$ be a random variable with $\mathbb{E}N_0 < \infty$ representing the initial number of targets in the mission space, and $\{T_i\}_{i\geq 1}$ be the sequence of random variables representing time between consecutive targets arrival, called interarrival times, where $T_i = 0$ for any $1 \leq i \leq N_0$, if $N_0 > 0$, and $\{T_i\}_{i>N_0}$ be i.i.d. non-negative random variables independent of N_0. The arrival time of i^{th} target can then be defined for any $i \in \mathbb{N}$ as following

$$\check{\tau}_i = \begin{cases} 0, & i \leq N_0, \\ \sum_{j=N_0+1}^{i} T_j, & i > N_0, \end{cases} \tag{3.7}$$

Let $\check{\mathcal{I}}_{\mathcal{T}}(t)$ denote the set of indices of targets arrived up to time moment t, i.e.

$$\check{\mathcal{I}}_{\mathcal{T}}(t) := \{i \in \mathbb{N} \mid \check{\tau}_i \leq t\}. \tag{3.8}$$

Let also $\{Y_i\}_{i\geq 1}$ be a sequence of i.i.d. random vectors in $\mathbb{R}^d$, independent of N_0 and $\{T_i\}_{i>N_0}$, with probability density function ϕ, a compact support absolutely continuous spatial distribution, such that $\text{supp}(\phi) \subseteq \mathcal{M}$. Having these all, one can say the i^{th} target arrives in the mission space at time $\check{\tau}_i$ and at point Y_i, for any $i \in \mathbb{N}$.

Definition 4. *Given a positive scalar $s_i, i \in \mathbb{N}$, the vehicle is said to visit the i^{th} target at time t, if $\|\mathrm{x}(t) - \mathrm{y}_i(t)\| \leq s_i$.*

Remark 3. *The scalar s_i in Definition 4 is introduced due to practical considerations regarding the physical size of the target. More precisely, while the dynamic equation of each target is expressed as a point mass, the scalar s_i is used to account for the size of the i^{th} target. For example, if the i^{th} target has a spherical shape of radius r_i and also the vehicle has a spherical shape with radius r, then $s_i := r_i + r$.*

A *task* is defined for every target, which is completed if the target is visited by the vehicle. Let $\hat{\tau}_i$ be the completion time of i^{th} task if the i^{th} target is visited in finite-time, and infinity if the i^{th} target is never visited by the vehicle. Similar to $\check{\mathcal{I}}_{\mathcal{T}}(t)$, one can define the set of indices of targets visited up to time t as

$$\hat{\mathcal{I}}_{\mathcal{T}}(t) := \{i \in \mathbb{N} \mid \hat{\tau}_i \leq t\}, \tag{3.9}$$

and the set of indices of targets arrived in the mission space but not visited up to time t by

$$\mathcal{I}_{\mathcal{T}}(t) := \check{\mathcal{I}}_{\mathcal{T}}(t) \backslash \hat{\mathcal{I}}_{\mathcal{T}}(t) = \{i \in \mathbb{N} \mid \check{\tau}_i \leq t < \hat{\tau}_i\}. \tag{3.10}$$

Let $N(t) = |\mathcal{I}_{\mathcal{T}}(t)|$, $\check{N}(t) = |\check{\mathcal{I}}_{\mathcal{T}}(t)|$ and $\hat{N}(t) = |\hat{\mathcal{I}}_{\mathcal{T}}(t)|$. It is to be noted that $\check{N}(t) = \hat{N}(t) + N(t)$ and $0 \leq N(t) \leq \check{N}(t)$. Note also that $\check{N}(t) - N_0$ is the counting process for the renewal process defined by $\{T_i\}_{i \geq N_0}$.

The trajectory of each target is a C^1 curve in the mission space $\mathcal{M}$, and is assumed to satisfy the geometric conditions given below.

Assumption 4. *For any $i \in \mathbb{N}$ and any $s \in [\check{\tau}_i, \hat{\tau}_i)$, it is assumed that*

- *(Global Geometric Condition) If $\mathrm{y}_i(\tau) \in \mathcal{M}$, then $\mathrm{y}_i(t) \in \mathcal{M}$ for all $t \in [\tau, \hat{\tau}_i)$.*

- *(Local Geometric Condition) There exist non-negative scalars v, B such that*

$$\left\| \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{y}_i(\tau) \right\| \leq v, \tag{3.11}$$

  and

$$\sup_{t \in [\tau, D]} \|\alpha_i(t, \tau)\| \leq B, \tag{3.12}$$

where $\alpha_i(t,\tau)$ *is a* C^1 *function satisfying the following equality*

$$\mathrm{y}_i(t) = \mathrm{y}_i(\tau) + \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(\tau)(t-\tau) + \frac{1}{2}\alpha_i(t,\tau)(t-\tau)^2. \tag{3.13}$$

The global geometric condition on targets' trajectories guarantees that once a target is detected in the mission space, it will remain inside it throughout the rest of the mission. This property is dependent on the targets' trajectories as well as the geometry of the mission space. In the particular, the global geometric condition is satisfied when $\mathcal{M} = \mathbb{R}^d$.

Regarding the local geometric condition, if $\mathrm{y}_i(t)$ be a C^2 function and there exist non-negative scalars v, B such that for any $i \in \mathbb{N}$ and $\tau \in [\check{\tau}_i, \hat{\tau}_i)$ the following conditions hold:

$$\left\|\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(\tau)\right\| \le v, \quad \left\|\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{y}_i(\tau)\right\| \le B. \tag{3.14}$$

then, from Taylor's theorem with mean-value form of the remainder [127], $\mathrm{y}_i(t)$ satisfies the local geometric condition in Assumption 4.

Assumption 5. *The position and velocity vectors are known at the beginning of each time horizon (i.e., at time* τ *in (3.13)).*

It follows from Assumption 5, that the positions of the existing targets can be estimated with sufficient accuracy at any future instant within the finite horizon. Let this estimate be denoted by $\hat{\mathrm{y}}_i(\cdot)$ for any $i \in \mathbb{N}$. Then

$$\hat{\mathrm{y}}_i(t) = \mathrm{y}_i(\tau) + \mathrm{v}_i(\tau)(t - \tau), \quad t \in [\tau, \hat{\tau}_i). \tag{3.15}$$

3.3 Receding Horizon Dynamic Decision-Making Scheme

In order to track the targets, a time-decreasing function called "reward" is assigned to every task which can be collected only if the target is visited. It is desired to plan vehicle's trajectory via dynamic decision-making such that the collected rewards are maximized. To this end, vehicle's trajectory is planned iteratively, where at the beginning of each iteration the heading and size of movement are calculated such that the estimate of the future collectible rewards is maximized.

3.3.1 Reward Functions

For any $i \in \mathbb{N}$, consider the reward function $R_i\rho_i$, where R_i denotes the initial reward for task i, the maximum reward considered for the target i, and $\rho_i : [0, \infty) \to [0, 1]$ is a decreasing function representing the rate of reward loss over time. Various candidates can be used for function ρ_i to model scheduling and time priorities. In particular, one can use the following discount function

$$\rho_i(t) = e^{-\gamma_i t}, \quad \forall\ i \in \mathbb{N} \tag{3.16}$$

where $\gamma_i \in \mathbb{R}_{>0}$ is a target-specific loss parameter reflecting different cases of interest.

One can now define the *total reward function* denoted by $\mathscr{R} : \mathbb{R} \to \mathbb{R}$ as

$$\mathscr{R}(t) := \sum_{i \in \mathbb{N}} R_i \rho_i(t) \mathbf{1}_{[\check{\tau}_i, \hat{\tau}_i)}(t), \tag{3.17}$$

representing the net reward available at time $t \in \mathbb{R}_{\geq 0}$.

3.3.2 Dynamic Assignment Structure

Given the positions of the vehicle and available targets in the mission space $\mathcal{M}$, it is desired to properly assign tasks to the vehicles. For any $i \in \mathbb{N}$, one can generalize the bivalent assignment, where target i is either (fully) assigned to the vehicle or not assigned to it at all, to define the grade of assignment denoted by $a_i \in [0, 1]$. More precisely, for any $t \in [0, \infty)$ and $i \in \mathbb{N}$, one can define functions

$$a_i : [0, \infty) \to [0, 1], \quad i \in \mathbb{N} \tag{3.18}$$

reflecting the grade of assignment of target i to the vehicle, for every $i \in \mathbb{N}$ at any time instant. The function $a_i(t)$ is, in fact, the level of interest of the vehicle in target i being assigned to it at time $t \in [0, \infty)$.

Since the assigning strategy described above is only for existing targets, if the i^{th} target has not arrived yet or is visited before, then $a_i(t) = 0$, i.e., for any $i \in \mathbb{N}$ and $t \in \mathbb{R}_{\geq 0}$, one have $a_i(t) = \mathbf{1}_{[\check{\tau}_i, \hat{\tau}_i)}(t) a_i(t)$. Also, since there is only one vehicle, it is expected that when there are some targets, the vehicle net assignment should be equal to one, i.e., at any time t for which $N(t) \neq 0$,

$$\sum_{i \in \mathcal{I}_{\mathcal{T}}(t)} a_i(t) = 1 \tag{3.19}$$

or equivalently

$$\sum_{i \in \mathcal{I}_{\mathcal{T}}(t)} a_i(t) = \sum_{i \in \mathbb{N}} a_i(t) \mathbf{1}_{[\check{\tau}_i, \hat{\tau}_i)}(t) = 1 - \delta_{0N(t)}, \tag{3.20}$$

where δ is the Kronecker delta function. For any $\mathcal{I} \subset \mathbb{N}$, define the set

$$\mathcal{A}_{\mathcal{I}} := \{(a_i)_{i \in \mathbb{N}} \in \mathbb{R}_{\geq 0}^{\mathbb{N}} \mid \sum_{i \in \mathbb{N}} a_i = 1 - \delta_{0|\mathcal{I}|}, a_i = \mathbf{1}_{\mathcal{I}}(i) a_i\}. \tag{3.21}$$

Then, equation (3.20) can be rewritten as

$$(a_i(t))_{i \in \mathbb{N}} \in \mathcal{A}_{\mathcal{I}(t)}. \tag{3.22}$$

Now, define the *total assigned reward function* $\mathfrak{R} : \mathbb{R} \to \mathbb{R}$ as

$$\mathfrak{R}(t) := \sum_{i \in \mathbb{N}} a_i(t) R_i \rho_i(t) \mathbf{1}_{[\check{\tau}_i, \hat{\tau}_i)}(t), \tag{3.23}$$

which represents the net assigned reward available at time $t \in \mathbb{R}_{\geq 0}$. One can easily verify that

$$\mathfrak{R}(t) = \sum_{i \in \mathcal{I}_{\mathcal{T}}(t)} a_i(t) R_i \rho_i(t). \tag{3.24}$$

The vehicle needs to dynamically perform the optimal assignment and trajectory planning such that the collected rewards are maximized.

3.3.3 Receding Horizon Trajectory Construction

It is desired to develop a receding horizon-based dynamic decision-making controller (RHDDC), which iteratively controls the headings of the vehicle, step size and optimal assignment such that the collected rewards are maximized. Let $\{t_k\}_{k=0}^{\infty} \in \mathbb{R}_{\geq 0}$ denote the time instants at which the RHDDC is applied. At any t_k, an optimization problem is solved, which provides an estimate of the collectible rewards in the future. The optimization problem is formulated based on the current positions of the vehicle and targets as well as the predicted future position of the targets. The solution of the problem provides the optimal control input $\mathrm{u}_k := \mathrm{u}(t_k)$ along with the optimal assignment $\{a_i^k = a_i(t_k)\}_{i \in \mathcal{I}_{\mathcal{T}}}$.

Let H_k be the planning horizon for RHDDC. For simplicity, the action horizon is chosen to be the same as the planning horizon and consequently, $t_{k+1} = t_k + H_k$. Under control input $\mathrm{u}(t_k)$ the planned position of the vehicle at time t_{k+1} is given by

$$\mathrm{x}(t_{k+1}) = \mathrm{x}(t_k) + \mathrm{u}(t_k)H_k.$$

with the current positions of the vehicle and targets, the earliest possible time that the vehicle can visit target i is predicted to be

$$\tau_i(\mathrm{u}^k, t_k) = (t_k + H_k) + \frac{\|\mathrm{x}(t_k + H_k) - \hat{\mathrm{y}}_i(t_k + H_k)\|}{V_{\max} + v}.$$

$\tau_i(\mathrm{u}^k, t_k)$ given above will be the exact visiting time of vehicle and target i if they move toward each other with maximum speed, and also the estimate $\hat{\mathrm{y}}_i(t_k + H_k)$ is exact. Under these conditions, if $a_i(t_{k+1})$ is the optimal assignment, it is expected to remain unchanged until the vehicle visits target i. This implies that

$$a_i(\tau_i(\mathrm{u}^k, t_k)) = a_i(t_{k+1}). \tag{3.25}$$

Similarly, one can estimate the maximum total reward that the vehicle is expected to collect at the time t_{k+1}. Denote this predicted expected reward by $\mathfrak{R}^{k+1}(\mathrm{u}_k, t_k)$. Let $\tilde{\rho}_i(\mathrm{u}_k, t_k) = \rho_i[\tau_i(\mathrm{u}_k, t_k)]$ and $\tilde{a}_i(\mathrm{u}^k, t_k) = a_i(\mathrm{u}_k, t_k)$. It follows the definition of

$\mathfrak{R}^{k+1}(\mathrm{u}_k, t_k)$ that

$$\mathfrak{R}^{k+1}(\mathrm{u}_k, t_k) = \sum_{i \in \mathcal{I}_{\mathcal{T}}(t_k)} R_i \tilde{\rho}_i(\mathrm{u}_k, t_k) \tilde{a}_i(\mathrm{u}_k, t_k). \tag{3.26}$$

Note that $\mathfrak{R}^k(\mathrm{u}_k, t_k)$ represents an estimation at the current time t_k of the total reward that the vehicle can expect at the next time instant t_{k+1} to collect throughout the mission.

One can now present the optimization problem at the k^{th} iteration P^k, as follows:

$$\begin{aligned} \max \quad & \mathfrak{R}^{k+1}(\mathrm{u}_k, t_k) \\ \text{s.}t. \quad & \tilde{\mathrm{a}}(\mathrm{u}_k, t_k) \in \mathcal{A}^k, \\ & \mathrm{u}_k \in \mathcal{U}^k. \end{aligned} \tag{3.27}$$

where $\mathcal{A}^k = \mathcal{A}_{\mathcal{I}_{\mathcal{T}}(t_k)}$ and $\mathcal{U}^k = \{\mathrm{u} \in \mathbb{R}^d; \mathrm{u} \in \mathbb{R}^d, \|\mathrm{u}\| \leq V_{\max}\}$ is the set of *admissible heading control*.

For convenience of notation, $\mathrm{x}(t_k)$, $\mathrm{y}_i(t_k)$ and $\hat{\mathrm{y}}_i(t_k)$ will hereafter be denoted by x^k, y_i^k and $\hat{\mathrm{y}}_i^k$, respectively, for any $i \in \mathcal{I}_{\mathcal{T}}$. The corresponding vectors are represented by y^k and $\hat{\mathrm{y}}^k$, accordingly.

3.4 Trajectory Analysis in Receding Horizon Control

In this section, theoretical analysis on the vehicle's trajectory in the mission space is presented. It is assumed that the mission space $\mathcal{M}$ is compact with diameter $\text{diam}(\mathcal{M})$.

Definition 5. *The trajectory* $\mathrm{x}(t)$ *in* $\mathcal{M}$ *is called a stationary trajectory if for all* $i \in \mathbb{N}$*, the* i^{th} *target hitting time or* i^{th} *task completion time are almost sure finite, i.e. one has* $\mathbb{P}(\check{\tau}_i < \infty) = 1$*.*

In order to investigate the behavior of the system and showing the stationarity of vehicle's trajectory, one needs to analyze the optimization problem presented in (3.27), and also the asymptotic behavior of solutions sequence resulted from (3.27), given a stochastic process which models targets arrivals. To this ends, some lemmas and theorems are presented in the sequel.

3.4.1 Equivalent Optimization Problems

In order to present the stationarity analysis of the vehicles trajectory, an equivalency theorem for the optimization problem (3.27) is required first.

Lemma 2. *Let* Ω *be a compact subset of* $\mathbb{R}^d$ *with non-empty interior. Let also* $f_i : \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ *be a continuous function, for any* $i \in \mathbb{N}_n$ *and* $\mathcal{A} = \{\mathrm{a} \in \mathbb{R}^n \mid \mathrm{a} \geq 0, \mathrm{a}^{\mathsf{T}}\mathbf{1} = 1\}$*.*

Define the function $g_1 : \Omega \times \mathcal{A} \to \mathbb{R}$ *as*

$$g_1(\mathrm{x}, \mathrm{a}) = \sum_{i=1}^{n} a_i e^{-f_i(\mathrm{x})}, \tag{3.28}$$

and the function $g_2 : \Omega \times \mathcal{A} \to \mathbb{R}$ *as*

$$g_2(\mathrm{x}, \mathrm{a}) = \sum_{i=1}^{n} a_i f_i(\mathrm{x}). \tag{3.29}$$

Then,

$$\underset{(\mathrm{x},\mathrm{a})\in\Omega\times\mathcal{A}}{\operatorname{argmax}}\ g_1(\mathrm{x}, \mathrm{a}) = \underset{(\mathrm{x},\mathrm{a})\in\Omega\times\mathcal{A}}{\operatorname{argmin}}\ g_2(\mathrm{x}, \mathrm{a}). \tag{3.30}$$

Moreover, for any $(\mathrm{x}^*, \mathrm{a}^*)$ *in the aforementioned sets, there exists some* $i \in \mathbb{N}_n$ *such that* $g_1(\mathrm{x}^*, \mathrm{a}^*) = g_1(\mathrm{x}^*, \mathrm{e}_i)$, $g_2(\mathrm{x}^*, \mathrm{a}^*) = g_2(\mathrm{x}^*, \mathrm{e}_i)$ *and* $(\mathrm{x}^*, \mathrm{e}_i)$ *also belongs to these sets.*

Proof. Since $\mathcal{A}$ is a closed and bounded subset of $\mathbb{R}^d$, it is a compact set. Also, Ω is a compact set, and hence $\Omega \times \mathcal{A}$ is a compact set. Define $\mathcal{E}_1 = \operatorname{argmax}_{(\mathrm{x},\mathrm{a})\in\Omega\times\mathcal{A}} g_1(\mathrm{x}, \mathrm{a})$ and $\mathcal{E}_2 = \operatorname{argmin}_{(\mathrm{x},\mathrm{a})\in\Omega\times\mathcal{A}} g_2(\mathrm{x}, \mathrm{a})$. Then, from the continuity of g_1 and g_2, and also the compactness of $\Omega \times \mathcal{A}$, it follows that $\mathcal{E}_1$ and $\mathcal{E}_2$ are nonempty sets. In order to prove that $\mathcal{E}_1 = \mathcal{E}_2$, it is suffices to show that $\mathcal{E}_1 \supseteq \mathcal{E}_2$ and $\mathcal{E}_1 \subseteq \mathcal{E}_2$.

Before proceeding with the proof, define $f(\mathrm{x}) = \min_{i\in\mathbb{N}_d} f_i(\mathrm{x})$, $I(x) = \operatorname{argmin}_{i\in\mathbb{N}_d} f_i(\mathrm{x})$ and $i(\mathrm{x}) = \min I(\mathrm{x})$. Thus, for any $\mathrm{x} \in \Omega$ and $i \in I(\mathrm{x})$, one can conclude that

$f(\mathrm{x}) = f_i(\mathrm{x})$, and also,

$$\max_{\mathrm{a}\in\mathcal{A}} g_1(\mathrm{x},\mathrm{a}) = \max_{\mathrm{a}\in\mathcal{A}} \sum_{i=1}^{n} a_i e^{-f_i(\mathrm{x})} = e^{-f(\mathrm{x})} \tag{3.31}$$

and

$$\min_{\mathrm{a}\in\mathcal{A}} g_2(\mathrm{x},\mathrm{a}) = \min_{\mathrm{a}\in\mathcal{A}} \sum_{i=1}^{n} a_i f_i(\mathrm{x}) = f(\mathrm{x}). \tag{3.32}$$

Now, let $(\mathrm{x}^*, \mathrm{a}^*) \in \mathcal{E}_1$. Therefore, for any $(\mathrm{x}, \mathrm{a}) \in \Omega \times \mathcal{A}$, one has $g_1(\mathrm{x}^*, \mathrm{a}^*) \geq g_1(\mathrm{x}, \mathrm{a})$. On the other hand

$$g_1(\mathrm{x}^*, \mathrm{a}^*) = \max_{\mathrm{a}\in\mathcal{A}} \sum_{i=1}^{n} a_i e^{-f_i(\mathrm{x}^*)} = e^{-f(\mathrm{x}^*)}. \tag{3.33}$$

Thus, for any $i \in \mathbb{N}_n$ and $\mathrm{x} \in \Omega$, it follows that $e^{-f(\mathrm{x}^*)} \geq g_1(\mathrm{x}, \mathrm{e}_i) = e^{-f_i(\mathrm{x})}$, i.e., $\mathrm{x}^* \in \mathrm{argmin}_{\mathrm{x}\in\Omega} f(\mathrm{x})$. Also, for any $i \in \mathbb{N}_d \backslash I(\mathrm{x}^*)$, $a_i^* = 0$. Let $\mathrm{a} = \mathrm{a}^* - a_i^* \mathrm{e}_i + a_i^* \mathrm{e}_{i(\mathrm{x}^*)}$, to obtain $\mathrm{a} \in \mathcal{A}$ and

$$0 \geq g_1(\mathrm{x}^*, \mathrm{a}) - g_1(\mathrm{x}^*, \mathrm{a}^*) = a_i^* (e^{-f_{i(\mathrm{x}^*)}(\mathrm{x}^*)} - e^{-f_i(\mathrm{x}^*)}). \tag{3.34}$$

From the definition of $i(\mathrm{x}^*)$, it results that $f_{i(\mathrm{x}^*)}(\mathrm{x}^*) < f_i(\mathrm{x}^*)$, and consequently $e^{-f_{i(\mathrm{x}^*)}(\mathrm{x}^*)} > e^{-f_i(\mathrm{x}^*)}$. Therefore, the equation (3.34), yields that $a_i = 0$. As a result, for any $(\mathrm{x}, \mathrm{a}) \in$

$\Omega \times \mathcal{A}$, one obtains

$$g_2(\mathrm{x}^*, \mathrm{a}^*) = \sum_{i \in I(\mathrm{x}^*)} a_i^* f_i(\mathrm{x}^*) = f(\mathrm{x}^*) \tag{3.35}$$

and also

$$f(\mathrm{x}^*) = \sum_{i \in I(\mathrm{x}^*)} a_i f_i(\mathrm{x}^*) \leq g_2(\mathrm{x}, \mathrm{a}). \tag{3.36}$$

Thus, for any $(\mathrm{x}, \mathrm{a}) \in \Omega \times \mathcal{A}$, one has $g_2(\mathrm{x}^*, \mathrm{a}^*) \leq g_2(\mathrm{x}, \mathrm{a})$, and consequently $(\mathrm{x}^*, \mathrm{a}^*) \in \mathcal{E}_2$, and as a result $\mathcal{E}_1 \subseteq \mathcal{E}_2$.

Using a similar argument, one can show that $\mathcal{E}_1 \supseteq \mathcal{E}_2$, which yields $\mathcal{E}_1 = \mathcal{E}_2$.

Let $\mathcal{E} = \mathcal{E}_1$, which also means $\mathcal{E} = \mathcal{E}_2$. From the definition of $i(\mathrm{x})$ and also equations (3.31) and (3.32), for any $(\mathrm{x}^*, \mathrm{a}^*) \in \mathcal{E}$,

$$g_1(\mathrm{x}^*, \mathrm{a}^*) = e^{-f(\mathrm{x}^*)} = g_1(\mathrm{x}^*, \mathrm{e}_{i(\mathrm{x}^*)}) \tag{3.37}$$

and

$$g_2(\mathrm{x}^*, \mathrm{a}^*) = f(\mathrm{x}^*) = g_2(\mathrm{x}^*, \mathrm{e}_{i(\mathrm{x}^*)}). \tag{3.38}$$

This completes the proof. □

Corollary 2. *Let r be a real positive scalar, $\mathcal{M} \subseteq \mathbb{R}^d$ be a compact set with non-empty interior, w be a vector in $\mathcal{M}$, and $\{\mathrm{y}_i\}_{i \in \mathbb{N}_n}$ be n arbitrary points in $\mathcal{M}$. Then, for any*

$i \in \mathbb{N}_n$ and any choice of $\gamma_i \in \mathbb{R}_{>0}$, the following optimization problems

$$\begin{aligned} \max \quad & \textstyle\sum_{i=1}^n a_i e^{-\gamma_i \|\mathrm{x}-\mathrm{y}_i\|} \\ \mathrm{s.}t. \quad & \textstyle\sum_{i=1}^n a_i = 1, \\ & a_i \geq 0, \quad \forall i \in \mathbb{N}_n \\ & \|\mathrm{x} - \mathrm{w}\| \leq r, \\ & \mathrm{x} \in \mathcal{M}, \end{aligned} \tag{3.39}$$

and

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^n a_i \gamma_i \|\mathrm{x}-\mathrm{y}_i\| \\ \mathrm{s.}t. \quad & \textstyle\sum_{i=1}^n a_i = 1, \\ & a_i \geq 0, \quad \forall i \in \mathbb{N}_n \\ & \|\mathrm{x} - \mathrm{w}\| \leq r, \\ & \mathrm{x} \in \mathcal{M}, \end{aligned} \tag{3.40}$$

are equivalents, i.e. $(\mathrm{x}^*, \mathrm{a}^*)$ is a solution for problem (3.39) if and only if it is a solution for problem (3.40). Moreover, for any such solution $(\mathrm{x}^*, \mathrm{a}^*)$, there exists some $i \in \mathbb{N}_n$ such that $(\mathrm{x}^*, \mathrm{e}_i)$ is also a solution of problems (3.39) and (3.40) with the optimal values $e^{-\gamma_i \|\mathrm{x}^* - \mathrm{y}_i\|}$ and $\gamma_i \|\mathrm{x}^* - \mathrm{y}_i\|$, respectively.

Proof. Let $\Omega = \mathcal{M} \cap \{\mathrm{x} \in \mathbb{R}^d \mid \|\mathrm{x} - \mathrm{w}\| \leq r\}$, and for any $i \in \mathbb{N}_d$, define function $f_i : \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ as $f_i(\mathrm{x}) = \gamma_i \|\mathrm{x} - \mathrm{y}_i\|$. Then Ω is a compact set and f_i is a continuous function, for any $i \in \mathbb{N}_d$. The proof follows immediately from Lemma 2. $\square$

In the remainder of this chapter, for simplicity of the analysis, it is assumed that $R_i = R$ and $\gamma_i = \gamma$, for any $i \in \mathbb{N}$ and some positive scalars R and γ.

Theorem 3. *Let $J_{\mathrm{y}} : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ be defined as*

$$J_{\mathrm{y}}(\mathrm{x}, \mathrm{a}; \Gamma) = \sum_{i=1}^{d} a_i \gamma_i \|\mathrm{x} - \mathrm{y}_i\| \tag{3.41}$$

where $\Gamma = (\gamma_i)_{i\in\mathbb{N}_n}$ are n scalars in $\mathbb{R}_{\geq 0}$. Furthermore, $\{\mathrm{y}_i\}_{i\in\mathbb{N}_n}$ are n points in $\mathbb{R}^d$ and $\mathrm{y} = (\mathrm{y}_i)_{i\in\mathbb{N}_n}$. Then, if $\mathcal{I}_{\mathcal{T}}(t_k) \neq \emptyset$, the optimization problem P^k, presented in (3.27) is equivalent to

$$\begin{aligned} \min \quad & J_{\hat{\mathrm{y}}^k}(\mathrm{x}, \mathrm{a}; \Gamma^k) \\ \mathrm{s.}t. \quad & \mathrm{a} \in \mathcal{A}^k, \mathrm{x} \in \Omega^k. \end{aligned} \tag{3.42}$$

where $\hat{\mathrm{y}}^k$ is the position vector of the current targets, $\Gamma^k = \gamma/(V_{\max} + v)\mathbf{1}_{N(t_k)}$, and $\Omega^k = \mathscr{B}(\mathrm{x}^k, VH_k) \cap \mathcal{M}$.

Proof. Let control input $\mathrm{u}^k \in \mathcal{U}^k$ be applied over time interval $[t_k, t_k + H_k)$. It follows from the dynamics of vehicle that $\mathrm{x}^{k+1} = \mathrm{x}^k + \mathrm{u}^k H_k$. Therefore, since $\mathcal{U}^k = \{\mathrm{u} \in \mathbb{R}^d ; \mathrm{u} \in \mathbb{R}^d, \|\mathrm{u}\| \leq V_{\max}\}$, it results that

$$\mathrm{x}^{k+1} \in \Omega^k = \mathscr{B}(\mathrm{x}^k, VH_k) \cap \mathcal{M}. \tag{3.43}$$

From the above relation and the definition of $\mathcal{A}^k$, the optimization problems (3.27) and (3.42) have the same domain. Consider now the reward function $\mathfrak{R}^{k+1}(\mathrm{u}_k, t_k)$ in (3.27).

For any $i \in \mathbb{N}$, one has $\rho_i(t) = e^{-\gamma_i t}$ and

$$\tau_i(\mathrm{u}^k, t_k) = t_{k+1} + \frac{\|\mathrm{x}^{k+1} - \hat{\mathrm{y}}_i^{k+1}\|}{V_{\max} + v}.$$

Hence, it follows that

$$\tilde{\rho}_i(\mathrm{u}_k, t_k) = \gamma_t^k e^{-\gamma_v^k \|\mathrm{x}^{k+1} - \hat{\mathrm{y}}_i^{k+1}\|},$$

where $\gamma_t^k = e^{-\gamma t_{k+1}}$ and $\gamma_v^k = \gamma/(V_{\max} + v)$. Therefore, one obtain

$$\mathfrak{R}^{k+1}(\mathrm{u}_k, t_k) \;=\; R\; \gamma_t^k \sum_{i \in \mathcal{I}_{\mathcal{T}}(t_k)} \tilde{a}_i(\mathrm{u}_k, t_k) e^{-\gamma_v^k \|\mathrm{x}^{k+1} - \hat{\mathrm{y}}_i^{k+1}\|}.$$

Thus, using Corollary 2, the proof is concluded. □

3.4.2 Bounds for Sensitivity of Cost Function

The next two lemmas describe sensitivity of the function J, defined in (3.41), with respect to its arguments.

Lemma 3. *Let n be a natural number, $\mathcal{I}_1$ and $\mathcal{I}_2$ be two index set where $\mathcal{I}_1, \mathcal{I}_2 \subseteq \mathbb{N}_n$, Ω and $\mathcal{M}$ be compact sets in $\mathbb{R}^d$ such that $\Omega \subseteq \mathcal{M}$, $\{\mathrm{y}_i\}_{i \in \mathbb{N}_n}$ be n points in $\mathcal{M}$, the set $\mathcal{A}$ be defined as $\mathcal{A} = \{\mathrm{a} \in \mathbb{R}^n \mid \mathrm{a} \geq 0, \mathrm{a}^\mathsf{T}\mathbf{1} = 1\}$ and function $J_{\mathrm{y}} : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ be* *defined as in (3.41). For* $i \in \{1,2\}$*, define*

$$\begin{aligned} (\mathrm{x}_i^*, \mathrm{a}_i^*) \; = \; \operatorname{argmin} \quad & J_{\mathrm{y}}(\mathrm{x}, \mathrm{a}; \Gamma) \\ \mathrm{s.}t. \quad & \mathrm{x} \in \Omega, \\ & \mathrm{a} \in \mathcal{A} \cap (\cap_{j \in \mathbb{N}_n \setminus \mathcal{I}_i} \mathrm{e}_j^{\perp}). \end{aligned} \tag{3.44}$$

where $\Gamma = (\gamma_i)_{i \in \mathbb{N}_n}$ *are* n *scalars in* $\mathbb{R}_{\geq 0}$*. Then*

$$|J_{\mathrm{y}}(\mathrm{x}_2^*, \mathrm{a}_2^*) - J_{\mathrm{y}}(\mathrm{x}_1^*, \mathrm{a}_1^*)| \leq 2|\mathcal{I}_1 \Delta \mathcal{I}_2| \operatorname{diam}(\mathcal{M}) \max(\Gamma). \tag{3.45}$$

Proof. If $\mathcal{I}_1 = \mathcal{I}_2$, the proof is straightforward, since for both $i = 1$ and $i = 2$, one has the same problems in (3.47). Now consider the case $\mathcal{I}_1 \neq \mathcal{I}_2$. For any $i \in \{1,2\}$, let $\mathcal{A}_{\mathcal{I}_i} = \{\mathrm{a} \in \mathbb{R}^{|\mathcal{I}_i|} \mid \mathrm{a} \geq \mathbf{0}_{n_i}, \mathbf{1}_{|\mathcal{I}_i|}^{\intercal}\mathrm{a} = 1\}$. Subsequently, one can see that $\mathrm{a} \in \mathcal{A} \cap (\cap_{j \in \mathbb{N}_n \setminus \mathcal{I}_i} \mathrm{e}_j^{\perp})$ if and only if $\mathrm{a}_{\bullet \mathcal{I}_i} \in \mathcal{A}_{\mathcal{I}_i}$. Also, for any such a, one has

$$J_{\mathrm{y}}(\mathrm{x}, \mathrm{a}; \Gamma) = J_{\mathrm{y}_{\bullet \mathcal{I}_i}}(\mathrm{x}, \mathrm{a}_{\bullet \mathcal{I}_i}; \Gamma_{\bullet \mathcal{I}_i}). \tag{3.46}$$

Therefore, the optimization problem (3.47) is equivalent to

$$\begin{aligned} \operatorname{argmin} \quad & J_{\mathrm{y}_{\bullet \mathcal{I}_i}}(\mathrm{x}, \mathrm{a}_{\bullet \mathcal{I}_i}; \Gamma_{\bullet \mathcal{I}_i}) \\ \mathrm{s.}t. \quad & \mathrm{x} \in \Omega, \\ & \mathrm{a} \in \mathcal{A}_{\mathcal{I}_i}, \end{aligned} \tag{3.47}$$

with solution $(\mathrm{x}_i^*, \mathrm{a}_i^{*'})$. From Lemma 2, there exits standard vectors e_{j_i} in $\mathbb{R}^{|\mathcal{I}_i|}$ such that $j_i \in \mathcal{I}_i$ and

$$J_{\mathrm{y}_{\bullet\mathcal{I}_i}}(\mathrm{x}_i^*, \mathrm{a}_i^{*'}) = J_{\mathrm{y}_{\bullet\mathcal{I}_i}}(\mathrm{x}_i^*, \mathrm{e}_{j_i}) = \gamma_{j_i}\|\mathrm{x}^* - \mathrm{y}_{j_i}\|. \tag{3.48}$$

Therefore, from equation (3.46), non-negativity of γ_js and triangle inequality, it follows that

$$\begin{aligned} |J_{\mathrm{y}}(\mathrm{x}_2^*, \mathrm{a}_2^*) - J_{\mathrm{y}}(\mathrm{x}_1^*, \mathrm{a}_1^*)| &= |\gamma_{j_2}\|\mathrm{x}^* - \mathrm{y}_{j_2}\| - \gamma_{j_1}\|\mathrm{x}^* - \mathrm{y}_{j_1}\|| \\ &\leq \gamma_{j_2}\|\mathrm{x}^* - \mathrm{y}_{j_2}\| + \gamma_{j_1}\|\mathrm{x}^* - \mathrm{y}_{j_1}\| \end{aligned}$$

As $\|\mathrm{x}^* - \mathrm{y}_{j_2}\|, \|\mathrm{x}^* - \mathrm{y}_{j_1}\| \leq \mathrm{diam}(\mathcal{M})$ and $|\mathcal{I}_1 \Delta \mathcal{I}_2| \geq 1$, one obtains

$$\begin{aligned} |J_{\mathrm{y}}(\mathrm{x}_2^*, \mathrm{a}_2^*) - J_{\mathrm{y}}(\mathrm{x}_1^*, \mathrm{a}_1^*)| &\leq 2\,\mathrm{diam}(\Omega)\max(\Gamma) \\ &\leq 2|\mathcal{I}_1 \Delta \mathcal{I}_2|\,\mathrm{diam}(\mathcal{M})\max(\Gamma). \end{aligned}$$

This concludes the proof. □

Lemma 4. *Consider cost function* $J_{\mathrm{y}} : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ *defined in (3.41). Also, let* $\{\mathrm{y}_i\}_{i\in\mathbb{N}_n}$ *and* $\{\mathrm{y}'_i\}_{i\in\mathbb{N}_n}$ *be two sets of* n *points in* $\mathbb{R}^d$*, and* $\mathrm{y} = (\mathrm{y}_i)_{i\in\mathbb{N}_n}, \mathrm{y}' = (\mathrm{y}'_i)_{i\in\mathbb{N}_n}$*. Then, for any* $\mathrm{x} \in \mathcal{M}$ *and any* $\mathrm{a} \in \mathcal{A}$*, one has*

$$|J_{\mathrm{y}'}(\mathrm{x}, \mathrm{a}; \Gamma) - J_{\mathrm{y}}(\mathrm{x}, \mathrm{a}; \Gamma)| \leq \max(\Gamma) \max_{i\in\mathbb{N}_n} \|\mathrm{y}'_i - \mathrm{y}_i\|. \tag{3.49}$$

Proof. Due to definition of J_{y}, one has

$$J_{\mathrm{y}'}(\mathrm{x},\mathrm{a};\Gamma)-J_{\mathrm{y}}(\mathrm{x},\mathrm{a};\Gamma)=\sum_{i=1}^{n}a_i\gamma_i\|\mathrm{x}-\mathrm{y}_i\|-\sum_{i=1}^{n}a_i\gamma_i\|\mathrm{x}-\mathrm{y}_i'\|$$
$$=\sum_{i=1}^{n}a_i\gamma_i\left(\|\mathrm{x}-\mathrm{y}_i\|-\|\mathrm{x}-\mathrm{y}_i'\|\right).$$

Subsequently, from triangle inequality and non-negativity of a_is and γ_is, it follows that

$$|J_{\mathrm{y}'}(\mathrm{x},\mathrm{a};\Gamma)-J_{\mathrm{y}}(\mathrm{x},\mathrm{a};\Gamma)|\leq\sum_{i=1}^{n}a_i\gamma_i\ \big|\|\mathrm{x}-\mathrm{y}_i\|-\|\mathrm{x}-\mathrm{y}_i'\|\big|.$$

According to triangle equality and $\gamma_i \leq \max(\Gamma)$, for any i, one has

$$|J_{\mathrm{y}'}(\mathrm{x},\mathrm{a};\Gamma)-J_{\mathrm{y}}(\mathrm{x},\mathrm{a};\Gamma)|\leq\max(\Gamma)\sum_{i=1}^{n}a_i\ \|\mathrm{y}_i-\mathrm{y}_i'\|.$$

From $\sum_i a_i = 1$, it yields that

$$|J_{\mathrm{y}'}(\mathrm{x},\mathrm{a};\Gamma)-J_{\mathrm{y}}(\mathrm{x},\mathrm{a};\Gamma)|\leq\max(\Gamma)\max_{i\in\mathbb{N}_n}\|\mathrm{y}_i'-\mathrm{y}_i\|,$$

which proves the claim. □

3.4.3 Stationarity Analysis of Vehicle's Trajectory

The following Lemma is borrowed from Chapter 2.

Lemma 5. *Consider the vectors* $\mathrm{p}, q, v \in \mathbb{R}^d$ *and the set of non-negative real numbers* $V, H, B \in \mathbb{R}_{>0}$. *Assume that* $\|\mathrm{v}\| < V$, *and that* $\alpha : [0, H] \to \mathbb{R}^d$ *is a bounded vector-valued function defined over the interval* $[0, H]$ *with* $\max_{t\in[0,H]} \|\alpha(t)\| \le B$. *Define the set*

$$\Omega_{\mathrm{q},H} = \{(\mathrm{w}, t) \in \mathbb{R}^d \times \mathbb{R}; t \in [0, H], \|\mathrm{w} - \mathrm{q}\| \le Vt\}, \tag{3.50}$$

which is a convex compact subset of $\mathbb{R}^d$, *and let* $\mathrm{z} : [0, H] \to \mathbb{R}^d$ *be given by* $\mathrm{z}(t) = \mathrm{p} + \mathrm{v}t + \frac{1}{2}\alpha(t)t^2$. *Define also*

$$H < \min\{\frac{\|\mathrm{p} - \mathrm{q}\|}{V + \|\mathrm{v}\|}, \frac{V - \|\mathrm{v}\|}{B}\}, \tag{3.51}$$

and

$$(\mathrm{w}^*, t^*) = \underset{(\mathrm{w},t)\in\Omega_{\mathrm{q},H}}{\operatorname{argmin}} \frac{1}{2}\|\mathrm{w} - \mathrm{p} - \mathrm{v}t\|^2. \tag{3.52}$$

Then

i*)* $t^* = H$, $\|\mathrm{w}^* - \mathrm{q}\| = VH$*;*

ii*)*

$$\|\mathrm{p} - \mathrm{q}\| - \|\mathrm{w}^* - \mathrm{z}(H)\| \ge f(H) \tag{3.53}$$

where $f(h) = h(V - \|v\| - \frac{1}{2}Bh)$, *and*

iii) $\|\tilde{\mathrm{w}}^* - \mathrm{q}\| = VH$ *and* $\|\mathrm{p} - \mathrm{q}\| - \|\tilde{\mathrm{w}}^* - \mathrm{z}(H)\| \geq f(H)$, *where*

$$\tilde{\mathrm{w}}^* = \mathrm{argmin}_{\|\mathrm{w}-\mathrm{q}\| \leq VH} \|\mathrm{w} - \mathrm{p} - \mathrm{v}H\|. \tag{3.54}$$

Remark 4. *The function* $f(h) = h(V - \|\mathrm{v}\| - \frac{1}{2}Bh)$ *introduced in Lemma 5 is a concave quadratic function which is:* (i) *non-negative only in the interval* $I = [0, 2(V - \|\mathrm{v}\|)/B]$; (ii) *zero only at the endpoints of the interval;* (iii) *strictly increasing in the interval* $[0, (V - \|\mathrm{v}\|)/B]$, *and* (iv) *attains its maximum at the midpoint of the interval* I. *Therefore,* $f(H)$ *is positive, and if*

$$\|\mathrm{p} - \mathrm{p}\| \geq \frac{V^2 - \|\mathrm{v}\|^2}{B},$$

then the function f *takes its maximum value at* H, *i.e.*

$$f(h) \leq f(H) = \frac{(V - \|\mathrm{v}\|)^2}{2B}, \quad \forall h \in \mathbb{R}.$$

Following the equivalency results presented in Theorem 3, the stationarity analysis is presented now. Regarding $\{t_k\}_{k=0}^{\infty} \in \mathbb{R}_{\geq 0}$, time instants at which the RHDDC is applied, note that one has $t_{k+1} = t_k + H_k$. In order to construct this time instance sequence, let $\mathcal{I}^k = \mathcal{I}_{\mathcal{T}}(t_k)$ and

$$H_k = \min_{i \in \mathcal{I}^k} \{ \frac{||\mathrm{x}(t_k) - \mathrm{y}_i(t_k)|| - \frac{1}{2}s_i}{v + V}, \frac{V - v}{B} \}. \tag{3.55}$$

Note that as for any $t \in \mathbb{R}_{\geq 0}$, the set $\mathcal{I}_{\mathcal{T}}(t)$ is finite and also for any $i \in \mathcal{I}^k$, one has $||\mathrm{x}(t_k) - \mathrm{y}_i(t_k)|| > s_i$, then

$$H_k \geq \min\{ \frac{s}{2v + 2V}, \frac{V - v}{B} \}, \tag{3.56}$$

where $s = \min_{i \in \mathbb{N}} s_i > 0$. Therefore $t_k \uparrow \infty$ and RHDDC can operate at all times in a real-time fashion.. Now let

$$\mathfrak{J}^k := J_{\mathrm{y}^k_{\bullet \mathcal{I}^k}}(\mathrm{x}^k, \mathrm{a}^k) = \sum_{i \in \mathcal{I}_{\mathcal{T}}} a_i^k \gamma_i \| \mathrm{x}^k - \mathrm{y}_i^k \|, \tag{3.57}$$

where $(\mathrm{x}^k, \mathrm{a}^k)$ is derived from the (3.27) or equivalently (3.42) (see Theorem 3).

Lemma 6. *For any $k \in \mathbb{N}$, one has*

$$\begin{aligned} \mathfrak{J}^k - \mathfrak{J}^{k+1} \geq \frac{\gamma}{V_{\max} + v} [f(H_k) - & 2 \operatorname{diam}(\mathcal{M}) \Delta N^k \\ & - B H_{k-1}^2 - B H_k^2], \end{aligned} \tag{3.58}$$

where $\Delta N^k = \hat{N}(t_{k+1}) - \hat{N}(t_k) + \check{N}(t_{k+1}) - \check{N}(t_k)$.

Proof. Let $(\mathrm{x}^*, \mathrm{a}^*)$ be defined as follows

$$\begin{aligned}(\mathrm{x}^*, \mathrm{a}^*) \quad = \quad \operatorname{argmin} \quad & J_{\hat{\mathrm{y}}^{k+1}_{\bullet\mathcal{I}^k}}(\mathrm{x}, \mathrm{a}; \Gamma^k) \\ \text{s.}t. \quad & \mathrm{a} \in \mathcal{A}^k, \mathrm{x} \in \Omega^k.\end{aligned} \tag{3.59}$$

where $\hat{\mathrm{y}}^k$ is the vector of positions of current targets, $\Gamma^k = \gamma(V_{\max} + v)^{-1}\mathbf{1}_{N(t_k)}$ and $\Omega^k = \mathscr{B}(\mathrm{x}^k, VH_k) \cap \mathcal{M}$. Then from Lemma 4, one has

$$|J_{\hat{\mathrm{y}}^{k+1}_{\bullet\mathcal{I}^k}}(\mathrm{x}^*, \mathrm{a}^*; \Gamma^k) - J_{\mathrm{y}^{k+1}_{\bullet\mathcal{I}^k}}(\mathrm{x}^*, \mathrm{a}^*; \Gamma^k)| \leq \frac{\gamma B}{2V_{\max} + 2v} H_k^2. \tag{3.60}$$

Also, from the same Lemma, it follows that

$$|J_{\hat{\mathrm{y}}^{k}_{\bullet\mathcal{I}^k}}(\mathrm{x}^k, \mathrm{a}^k; \Gamma^k) - J_{\mathrm{y}^{k}_{\bullet\mathcal{I}^k}}(\mathrm{x}^k, \mathrm{a}^k; \Gamma^k)| \leq \frac{\gamma B}{2V_{\max} + 2v} H_{k-1}^2. \tag{3.61}$$

Moreover, Lemma 3 results in

$$|J_{\mathrm{y}^{k+1}_{\bullet\mathcal{I}^k}}(\mathrm{x}^*, \mathrm{a}^*; \Gamma^k) - J_{\mathrm{y}^{k+1}_{\bullet\mathcal{I}^{k+1}}}(\mathrm{x}^{k+1}, \mathrm{a}^{k+1}; \Gamma^{k+1})| \leq 2|\mathcal{I}^k \Delta \mathcal{I}^{k+1}| \frac{\gamma \operatorname{diam}(\mathcal{M})}{V_{\max} + v}. \tag{3.62}$$

Since $\Delta N^k = |\mathcal{I}^k \Delta \mathcal{I}^{k+1}|$, if one has

$$J_{\hat{\mathrm{y}}^{k}_{\bullet\mathcal{I}^k}}(\mathrm{x}^k, \mathrm{a}^k; \Gamma^k) - J_{\hat{\mathrm{y}}^{k+1}_{\bullet\mathcal{I}^k}}(\mathrm{x}^*, \mathrm{a}^*; \Gamma^k) \geq \frac{\gamma}{V_{\max} + v}[f(H_k) - \frac{1}{2}BH_{k-1}^2 - \frac{1}{2}BH_k^2], \tag{3.63}$$

then using triangle inequality and equations (3.60-3.63), one obtains

$$\mathfrak{J}^k - \mathfrak{J}^{k+1} \geq \frac{\gamma}{V_{\max} + v}[f(H_k) - 2\,\mathrm{diam}(\mathcal{M})\Delta N^k - BH_{k-1}^2 - BH_k^2], \tag{3.64}$$

which is the claim. Now, in order to bridge the gap, one has to prove (3.63). From Theorem 3, one can see that there exists $i^k \in \mathcal{I}^k$ such that

$$J_{\hat{\mathrm{y}}^k_{\bullet\mathcal{I}^k}}(\mathrm{x}^k, \mathrm{a}^k; \Gamma^k) = J_{\hat{\mathrm{y}}^k_{\bullet\mathcal{I}^k}}(\mathrm{x}^k, \mathrm{e}_{i^k}; \Gamma^k) = \frac{\gamma}{V_{\max} + v}\|\mathrm{x}^k - \hat{\mathrm{y}}^k_{i^k}\|. \tag{3.65}$$

Since $\|\hat{\mathrm{y}}^k_{i^k} - \mathrm{y}^k_{i^k}\| \leq \frac{1}{2}BH_{k-1}^2$, it follows that

$$J_{\hat{\mathrm{y}}^k_{\bullet\mathcal{I}^k}}(\mathrm{x}^k, \mathrm{a}^k; \Gamma^k) \geq \frac{\gamma}{V_{\max} + v}[\|\mathrm{x}^k - \mathrm{y}^k_{i^k}\| - \frac{1}{2}BH_{k-1}^2]. \tag{3.66}$$

Also, from equation (3.59) one can see

$$J_{\hat{\mathrm{y}}^{k+1}_{\bullet\mathcal{I}^k}}(\mathrm{x}^*, \mathrm{a}^*; \Gamma^k) \leq \frac{\gamma}{V_{\max} + v}\|\mathrm{w}^* - \hat{\mathrm{y}}^{k+1}_{i^k}\|, \tag{3.67}$$

where $\mathrm{w}^* = \mathrm{argmin}_{\mathrm{w} \in \Omega^k} \|\mathrm{w} - \hat{\mathrm{y}}^{k+1}_{i^k}\|$. From Lemma 5, it yields that

$$\|\mathrm{x}^k - \mathrm{y}^k_{i^k}\| - \|\mathrm{w}^* - \mathrm{y}^{k+1}_{i^k}\| \geq f(H_k). \tag{3.68}$$

Note that as $\|\hat{\mathrm{y}}_{i^k}^{k+1} - \mathrm{y}_{i^k}^{k+1}\| \leq \frac{1}{2} B H_{k-1}^2$, one has

$$\|\mathrm{x}^k - \mathrm{y}_{i^k}^k\| - \|\mathrm{w}^* - \hat{\mathrm{y}}_{i^k}^{k+1}\| \geq f(H_k) - \frac{1}{2} B H_{k-1}^2. \tag{3.69}$$

From the these, one can easily see that the inequality (3.63) holds. This concludes the proof. □

Lemma 7. *As $t \to \infty$, $\check{N}(t)/t$ converges $\mathbb{P}$-almost sure to $\lambda = 1/\mathbb{E}T$, i.e.*

$$\mathbb{P}(\lim_{t\to\infty} \frac{\check{N}(t)}{t} = \frac{1}{\mathbb{E}T}) = 1. \tag{3.70}$$

Proof. From $\mathbb{E}N_0 < \infty$, one can see that $\mathbb{P}(N_0 < \infty) = 1$. Subsequently, it follows that

$$\mathbb{P}(\lim_{t\to\infty} \frac{\check{N}_0}{t} = 0) = 1. \tag{3.71}$$

Also, as $\check{N}(t) - N_0$ is the counting process for the renewal process with parameter $1/\mathbb{E}T$, from Theorem 2, one can see that

$$\mathbb{P}(\lim_{t\to\infty} \frac{\check{N}(t) - N_0}{t} = \frac{1}{\mathbb{E}T}) = 1. \tag{3.72}$$

Subsequently, from equations (3.71) and (3.72), it follows that

$$\mathbb{P}(\lim_{t\to\infty} \frac{\check{N}(t)}{t} = \frac{1}{\mathbb{E}T}) = 1. \tag{3.73}$$

This concludes the proof. □

The main result of this work, stated in the next theorem, follows from Lemma 7 and the above discussion.

Theorem 4. *Consider the receding horizon problem presented in (3.27). Let*

$$\frac{1}{\mathbb{E}T} = \lambda < \frac{V - v}{4\,\mathrm{diam}(\mathcal{M})}, \tag{3.74}$$

and H_k sets due to

$$H_k = \min_{i \in \mathcal{I}^k} \{ \frac{||\mathrm{x}(t_k) - \mathrm{y}_i(t_k)|| - \frac{1}{2}s_i}{v + V}, \frac{V - v}{3B}, 2\frac{V - v - 4\,\mathrm{diam}(\mathcal{M})(\lambda + \epsilon)}{5B} \}, \tag{3.75}$$

where $0 < \epsilon$ is a scalar such that $\lambda + \epsilon < (V - v)/(4\,\mathrm{diam}(\mathcal{M}))$. Then with probability one there exist a time $t \in \mathbb{R}_{>0}$ such that $N(t) = 0$.

Proof. Assume that for any time $t \in \mathbb{R}_{>0}$, one has $N(t) \geq 1$. Then for all $k \in \mathbb{N}$, it follows that

$$\mathfrak{J}^k - \mathfrak{J}^{k+1} \geq \frac{\gamma}{V_{\max} + v}[f(H_k) - 2\,\mathrm{diam}(\mathcal{M})\Delta N^k - BH_{k-1}^2 - BH_k^2]. \tag{3.76}$$

Therefore, for any $K > 0$, letting $H_{-1} = 0$, it is concluded by induction that

$$\mathfrak{J}^0 - \mathfrak{J}^{K+1} \geq \frac{\gamma}{V_{\max} + v}[\sum_{k=0}^{K} f(H_k) - 2\,\mathrm{diam}(\mathcal{M})\sum_{k=0}^{K}\Delta N^k - \sum_{k=0}^{K} BH_{k-1}^2 - \sum_{k=0}^{K} BH_k^2]. \tag{3.77}$$

Hence, as $f(h) = h(V - v - \frac{1}{2}Bh)$ and $\Delta N^k = \hat{N}(t_{k+1}) - \hat{N}(t_k) + \check{N}(t_{k+1}) - \check{N}(t_k)$, one obtains

$$\mathfrak{J}^0 - \mathfrak{J}^{K+1} \geq \nu[\kappa \sum_{k=0}^{K} H_k - \check{N}(t_{K+1}) - 5\eta \sum_{k=0}^{K} H_k^2 + 2\eta H_K^2]. \tag{3.78}$$

where

$$\nu = \frac{4\gamma\,\mathrm{diam}(\mathcal{M})}{V_{\max} + v}, \quad \kappa = \frac{V - v}{4\,\mathrm{diam}(\mathcal{M})}, \quad \eta = \frac{B}{8\,\mathrm{diam}(\mathcal{M})}. \tag{3.79}$$

Therefore, it follows that

$$\mathfrak{J}^0 - \mathfrak{J}^{K+1} \geq \nu[\kappa \sum_{k=0}^{K} H_k - \check{N}(t_{K+1}) - 5\eta \sum_{k=0}^{K} H_k^2]. \tag{3.80}$$

Due to (3.75), one can see $\kappa - \lambda - \epsilon - 5\eta H_k \geq 0$ and $H_k > 0$, for any $k \in \mathbb{N}$. Subsequently, it follows that

$$\sum_{k=0}^{K} H_k(\kappa - \lambda - \epsilon - 5\eta H_k) \geq 0.$$

Since $t_{K+1} = \sum_{k=0}^{K} H_k$, one has

$$\frac{1}{\nu}\frac{\mathfrak{J}^0 - \mathfrak{J}^{K+1}}{t_{K+1}} \geq \epsilon + (\lambda - \frac{\check{N}(t_{K+1})}{t_{K+1}}). \tag{3.81}$$

From equation (3.57), it is concluded that $\|\mathfrak{J}^k\| \leq \max(\Gamma)\,\mathrm{diam}(\mathcal{M})$. Therefore, one has

$$\lim_{K\to\infty} \frac{1}{\nu}\frac{\mathfrak{J}^0 - \mathfrak{J}^{K+1}}{t_{K+1}} = 0, \tag{3.82}$$

Meanwhile, it follows from Lemma 7 that

$$\epsilon + (\lambda - \frac{\check{N}(t_{K+1})}{t_{K+1}}) \to_{K\to\infty} \epsilon > 0, \quad \mathbb{P}\text{-a.s.} \tag{3.83}$$

which says that the probability of the event that the inequality $N(t) \geq 1$ holds for any time $t \in \mathbb{R}_{>0}$ is zero. Hence, there exist almost surely a finite time $\tau \in \mathbb{R}_{>0}$ such that $N(\tau) = 0$. □

Definition 6. *For any $j \in \mathbb{N}$, let the j^{th} reseting time and j^{th} restarting time be random variables, denoted by τ^j and σ^j, respectively. Let also $\tau^0 = 0$ and $\sigma^0 = 0$, and define by $\tau^j = \inf\{t \in \mathbb{R}_{\geq 0} \mid N(t) = 0, t \geq \sigma^j\}$ and $\sigma^j = \inf\{t \in \mathbb{R}_{\geq 0} \mid N(t) \neq 0, t \geq \tau^{j-1}\}$, for any $j \in \mathbb{N}$.*

Remark 5. *Theorem 4 states that $\mathcal{P}(\tau^0 < \infty) = 1$, and it guarantees that after finite time the mission space resets to the state with no present target. This implies that*

$\mathcal{P}(\tau^j - \sigma^j < \infty) = 1$, *for any* $j \in \mathbb{Z}_{\geq 0}$.

Definition 7. *For any* $i \in \mathbb{N}$, *time window of the* i^{th} *task is a non-negative random variable denoted by* w_i, *and is defined as* $w_i = \hat{t}_i - \check{t}_i$.

Remark 6. *The time window of the* i^{th} *task represents the time interval between the* i^{th} *target arrival and being visited by the vehicle.*

Theorem 5. *Under the assumptions of Theorem 4, for any* $i \in \mathbb{N}$, *the time window of the* i^{th} *task has a finite value, i.e.* $\mathcal{P}(w_i < \infty) = 1$, *with probability one.*

Proof. For any $i \in \mathbb{N}$, due to Definition 6, there exits a unique $j \in \mathbb{Z}_{\geq 0}$ such that $\sigma^j \leq \check{t}_i \leq \tau^j$ and consequently, $\sigma^j \leq \hat{t}_i \leq \tau^j$. Thus, $w_i = \hat{t}_i - \check{t}_i \leq \tau^j - \sigma^j$, and from Theorem 4 and Remark 5, one can conclude that $\mathcal{P}(w_i < \infty) = 1$. □

3.5 Simulation Results

Simulations are presented in this section to verify the effectiveness of the proposed receding horizon dynamic decision-making scheme in target tracking.

Figure 3.1 illustrates the case where a vehicle (blue trajectory) with maximum velocity $V = 7$ is supposed to visit a sequence of four arriving targets (red trajectories) with $v = 5$ and $B = 1$ (see (3.11) and (3.12)). Here, the initial number of targets is two and $\lambda^{-1} = 10$. The vector of arrival times and vector of task completion times are $[\check{t}_1, \check{t}_2, \check{t}_3, \check{t}_4] = [0, 0, 22.4, 22.8]$ and $[\hat{t}_1, \hat{t}_2, \hat{t}_3, \hat{t}_4] = [7, 17, 28, 38.8]$, respectively. Thus, the

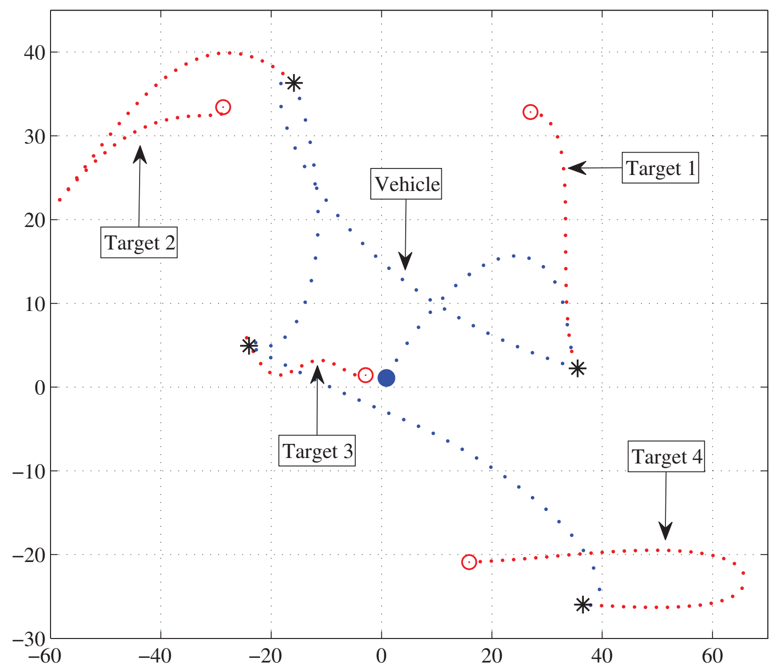

Figure 3.1: The vehicle's trajectory (the blue curve starting from the blue bullets in the origin) and targets' trajectory (the red curves, starting from the red circles). The positions where the vehicle visited the targets is shown by asterisks.

vector of time windows is $[w_1, w_2, w_3, w_4] = [7, 17, 5.6, 16]$. Note that in the time instants where there is no target in the mission space (e.g., $[\hat{t}_2, \check{t}_3] = [17, 22.4]$), the vehicle does not move.

Three scenarios are considered in Figure 3.2, representing infrequent target arrivals, medium rate of target arrivals, and frequent target arrivals case, corresponding to small, medium and large λ. The simulation parameters are $N_0 = 4, \mathcal{M} = [-90, 90]^2, V = 7, v = 2, B = 1$ and $r_i + r = 0.5$. Under the proposed control strategy, the results depicted in Figure 1 are obtained. Figure 3.2(a) provides the results for the case of infrequent targets, where λ^{-1} is assumed to be 10. The figure demonstrates that the number of

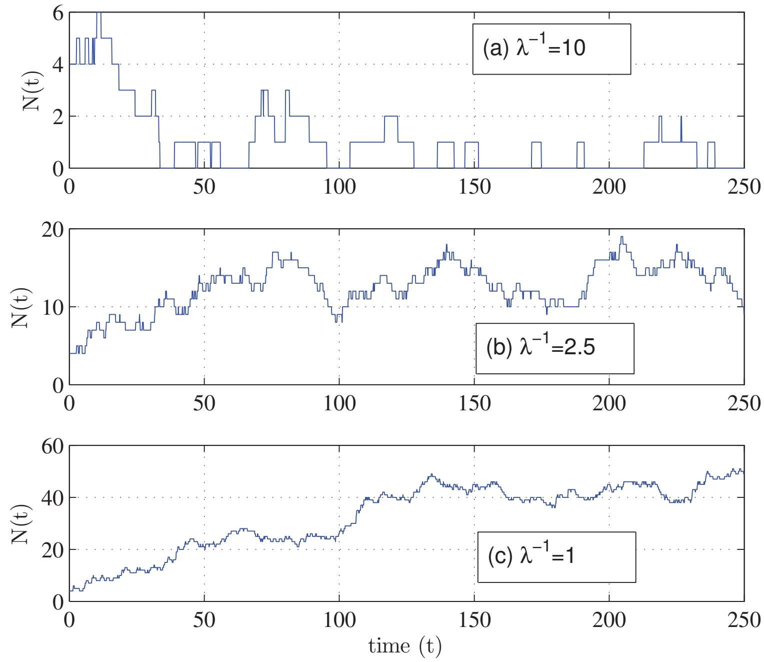

Figure 3.2: The number of targets in the mission space in the case of infrequent target arrivals (top figure), medium rate of target arrivals (middle figure), and frequent target arrivals (bottom figure).

remaining targets $N(t)$ in the mission space becomes equal to zero at some moment of time intervals throughout the operation of the system. This means that the vehicle can visit every arriving target.

Figure 3.2(b) presents the case of medium rate of target arrivals,, and in particular λ^{-1} in this case is chosen to be 2.5. The figure shows that in this scenario there is a balance in the operation of the system, in terms of target arrivals and vehicle's ability to visit them.

Figure 3.2(c) gives the results for the case of frequent target arrivals, where λ^{-1} is assumed to be equal to 1. As can be observed from this figure, the rate of target arrivals is too high for the vehicle to visit them, and hence the number of targets increases continuously.

Chapter 4

Cooperative Control for Multi-Target Interception with Sensing and Communication Limitations: A Game-Theoretic Approach

In this chapter, the problem of multi-vehicle cooperative interception of moving objects with *a priori* unknown arrival times, trajectories and dynamics is investigated. The vehicles are assumed to have limited sensing and communication ranges. Therefore, centralized approaches are not feasible, specially when there are a large number of vehicles and targets. A game-theoretic cooperative receding horizon controller is proposed, which predicts the future positions of targets with limited information. It uses a reward allocation policy for accomplishing the target interception task. To learn the optimal strategy in the resulting potential game, the generalized regret monitoring is used and its effectiveness is demonstrated by simulation.

4.1 Background

Throughout the chapter, $\mathbb{N}, \mathbb{R}, \mathbb{R}_{\geq 0}$ denote the set of natural numbers, real numbers, and non-negative real numbers, respectively. Also, the set of natural numbers less than or equal to k is denoted by $\mathbb{N}_k$. For a given set A, and some subset of it B, the indicator function of B is denoted by $\mathbf{1}_B$, which is a binary function from A to $\{0, 1\}$ as follows:

$$\mathbf{1}_B(x) = \begin{cases} 1, & x \in B \\ 0, & x \notin B \end{cases} \tag{4.1}$$

For any index set $\mathcal{I}$, the $A^{\mathcal{I}}$ represents the set of points like $(a_i)_{i\in\mathcal{I}}$ where each of its entries belongs A. In the case that $\mathcal{I}$ is the set $\mathbb{N}_n$, the $A^{\mathcal{I}}$ is simply shown by A^n. For any point in $\mathrm{a} \in A^{\mathcal{I}}$, the $\mathrm{a}_{\bullet\mathcal{J}}$, where $\mathcal{J}$ be a non-empty subset of $\mathcal{I}$, represents a point in $A^{\mathcal{J}}$, obtained by eliminating the entries with indices not listed in $\mathcal{J}$. The d dimensional Euclidean space is denoted by $\mathbb{R}^d$. Also, an all-zero vector and an all-one vector in $\mathbb{R}^d$

are respectively represented by $\mathbf{0}_d$ and $\mathbf{1}_d$. For any vectors a and b in $\mathbb{R}^d$, the inequality $\mathrm{a} \geq \mathrm{b}$ means that all entries of $\mathrm{a} - \mathrm{b}$ are non-negative. For any point $\mathrm{x} \in \mathbb{R}^d$ and any scalar $r \in \mathbb{R}_{\geq 0}$, a closed ball of radius r centered at x is denoted by $\mathscr{B}(\mathrm{x}, r)$, and is defined as $\{\mathrm{y} \in \mathbb{R}^d \mid \|\mathrm{x} - \mathrm{y}\| \leq r\}$, where $\|.\|$ represents the Euclidean norm. Denote by $C^p_{\mathbb{R}_{\geq 0}}(\mathbb{R}^d)$ the set of piecewise continuous functions defined over $\mathbb{R}_{\geq 0}$ and taking values in $\mathbb{R}^d$. A *bipartite graph* $\mathcal{G} = (\mathcal{U} \cup \mathcal{V}, \mathcal{E})$ is a graph whose vertex set is the union of two disjoint subsets $\mathcal{U}$ and $\mathcal{V}$, with no pair of adjacent vertices in each one. The *biadjacency matrix* of a bipartite graph $\mathcal{G}(\mathcal{U} \cup \mathcal{V}, \mathcal{E})$ is defined as a $|\mathcal{U}|$ by $|\mathcal{V}|$ matrix $\mathbf{B} = (b_{ij})$ of binary entries, where the (i, j) element is one if the i^{th} vertex in $\mathcal{U}$ is adjacent to the j^{th} vertex in $\mathcal{V}$, and zero otherwise.

A game of n players is represented by $(\mathbb{N}_n, \times_{i\in\mathbb{N}_n}\mathcal{A}_i, \{U_i\}_{i\in\mathbb{N}_n})$, where $\mathbb{N}_n$ is the players' index set, and for any $i \in \mathbb{N}_n$, $\mathcal{A}_i$ is the action set, $\bigtimes_{i\in\mathbb{N}_n} \mathcal{A}_i$ is the set of action profiles, and $U_i : \bigtimes_{i\in\mathbb{N}_n} \mathcal{A}_i \to \mathbb{R}$ is the utility function. For any $i \in \mathbb{N}_n$ and any action profile $(a_j)_{j\in\mathbb{N}_n} \in \times_{j\in\mathbb{N}_n}\mathcal{A}_j$, let a_{-i} and (a_i, a_{-i}) denote $(a_j)_{j\neq i}$ and $(a_j)_{j\in\mathbb{N}_n}$, respectively.

Definition 8 ([132]). *The game* $(\mathbb{N}_n, \bigtimes_{i\in\mathbb{N}_n}\mathcal{A}_i, \{U_i\}_{i\in\mathbb{N}_n})$ *is called a* potential game, *if there exists a function* $\phi : \times_{i\in\mathbb{N}_n}\mathcal{A}_i \to \mathbb{R}$, *called* potential function, *such that for any* $i \in \mathbb{N}_n$, *any actions* $a'_i, a''_i \in \mathcal{A}_i$ *and any* $\mathrm{a}_{-i} \in \times_{j\neq i}\mathcal{A}_j$, *the following relation holds*

$$U_i(a'_i, \mathrm{a}_{-i}) - U_i(a''_i, \mathrm{a}_{-i}) = \phi(a'_i, \mathrm{a}_{-i}) - \phi(a''_i, \mathrm{a}_{-i}).$$

4.2 Problem Formulation

Define the *mission space* as a closed convex subset of $\mathbb{R}^d$, and denote it by $\mathcal{M}$. Consider a finite number of objects, referred to as *targets*, arriving in the mission space sequentially. One can specify the targets with respect to their arrival order by indices in $\mathcal{I}_{\mathcal{T}} = \mathbb{N}_n$, where n is the number of targets. Without loss of generality, it can be assumed that the mission starts at time $t = 0$, where $n_0 \in \{0\} \cup \mathbb{N}_{|\mathcal{I}_{\mathcal{T}}|}$ is the initial number of targets in the mission space. Let T_1 denote the arrival time of the first target, and set the finite sequence of non-negative real scalars $\{T_i\}_{i=2}^{|\mathcal{I}_{\mathcal{T}}|}$ as the time interval between the arrival times of consecutive targets $i-1$ and i, for any $i \in \mathbb{N}_{|\mathcal{I}_{\mathcal{T}}|}$. From the above definition, the arrival time of target i, denoted by $\check{\tau}_i$, is $\sum_{j=1}^{i} T_j$ for any $i \in \mathcal{I}_{\mathcal{T}}$. Also, for any $t \in \mathbb{R}_{\geq 0}$, denote by $\check{\mathcal{I}}_{\mathcal{T}}(t)$ the set of indices of targets arrived up to time t, i.e.

$$\check{\mathcal{I}}_{\mathcal{T}}(t) := \{i \in \mathcal{I}_{\mathcal{T}} \ ; \ \check{\tau}_i \leq t\}. \tag{4.2}$$

For any $i \in \mathcal{I}_{\mathcal{T}}$, let $\mathrm{y}_i \in \mathcal{M}$ be the initial position of target i as it arrives in the mission space. Thus, $\{\mathrm{y}_i\}_{i\in\mathcal{I}_{\mathcal{T}}}$ is the finite sequence of the initial positions of targets. Also, since the targets are assumed to be a set of moving objects in the mission space, by a slight abuse of notation, one can represent by $\mathrm{y}_i(\cdot)$ the trajectory of the i^{th} target after its arrival, for any $i \in \mathcal{I}_{\mathcal{T}}$ (note that $\mathrm{y}_i = \mathrm{y}_i(\check{\tau}_i)$). The arrival times are not known *a priori*, and no information is available about the target trajectories. In other words, $\check{\tau}_i$ and $\mathrm{y}_i(\cdot)$ are not known at any $t < \check{\tau}_i$.

In addition to the targets, there are a finite number of vehicles in $\mathcal{M}$ whose indices belong to the set $\mathcal{I}_{\mathcal{V}} = \mathbb{N}_m$, where m is the number of vehicles. For any $j \in \mathcal{I}_{\mathcal{V}}$, denote by $\mathrm{x}_j(t)$ the position of vehicle j in the mission space at time t. Also, let the dynamics of $\mathrm{x}_j(t)$ be described by

$$\dot{\mathrm{x}}_j(t) = \mathrm{u}_j(t). \tag{4.3}$$

The input vector u_j in the above equation belongs to the set of *admissible controls* $\mathscr{U}_{u_{\max}}$, defined as

$$\mathscr{U}_{u_{\max}} = \{\mathrm{u} : \mathbb{R}_{\geq 0} \to \mathbb{R}^d \ ; \ \mathrm{u} \in C^p_{\mathbb{R}_{\geq 0}}(\mathbb{R}^d), \|\mathrm{u}\|_{\sup} \leq u_{\max}\}, \tag{4.4}$$

where $C^p_{\mathbb{R}_{\geq 0}}(\mathbb{R}^d)$ is the set of piecewise-continuous functions defined over $\mathbb{R}_{\geq 0}$, taking values in $\mathbb{R}^d$. Note that here $u_{\max}$ is a positive real scalar. It can be seen that for some piecewise continuous functions u_j and d_j, one has $\mathrm{u}_j(t) = u_j(t)\mathrm{d}_j(t)$, where $u_j(t) \in [0, u_{\max}]$ is the control input for the magnitude of the velocity vector of vehicle j, and $\mathrm{d}_j(t) \in \mathbb{S}^{d-1} = \{\mathrm{d} \in \mathbb{R}^d; \|\mathrm{d}\| = 1\}$ is the control input for its direction, for any $j \in \mathcal{I}_{\mathcal{V}}$ at any time $t \in \mathbb{R}_{\geq 0}$.

For any $j \in \mathcal{I}_{\mathcal{V}}$, define the *sensing region* of vehicle j at time $t \in \mathbb{R}_{\geq 0}$ as

$$\mathscr{S}_j(t) = \{\mathrm{x} \in \mathbb{R}^d \ ; \ \|\mathrm{x} - \mathrm{x}_j(t)\| \leq r_{sj}\} = \mathscr{B}(\mathrm{x}_j(t); r_{sj}), \tag{4.5}$$

where $r_{sj} \in \bar{\mathbb{R}}_{\geq 0}$ is a scalar representing the *sensing radius* of the vehicle. Similarly, for any $j \in \mathcal{I}_{\mathcal{V}}$ and at any time $t \in \mathbb{R}_{\geq 0}$, define the *communication region* of vehicle j as

$$\mathscr{C}_j(t) = \mathscr{B}(\mathrm{x}_j(t); r_c), \tag{4.6}$$

where $r_c \in \bar{\mathbb{R}}_{\geq 0}$ is a scalar representing the *communication radius* of the vehicle. For any $i \in \mathcal{I}_{\mathcal{T}}$ and $j \in \mathcal{I}_{\mathcal{V}}$, vehicle j is capable of sensing target i at time t, if the target is in the sensing region of the vehicle (i.e. $\mathrm{y}_i(t) \in \mathscr{S}_j(t)$). Also, for any pair of vehicles $j_1, j_2 \in \mathcal{I}_{\mathcal{V}}$ at any time t, vehicle j_2 can receive information sent by vehicle j_1, if vehicle j_2 is in the communication region of vehicle j_1 (i.e. $\mathrm{x}_{j_2}(t) \in \mathscr{C}_{j_1}(t)$).

Remark 7. *Note that for any $j_1, j_2 \in \mathcal{I}_{\mathcal{V}}$ at any time t, $\mathrm{x}_{j_2}(t) \in \mathscr{C}_{j_1}(t)$ if and only if $\mathrm{x}_{j_1}(t) \in \mathscr{C}_{j_2}(t)$. This means that the sensing network is symmetrical.*

Definition 9. *For any $i \in \mathcal{I}_{\mathcal{T}}$ and $j \in \mathcal{I}_{\mathcal{V}}$, vehicle j is said to* visit *target i at time t, if $\|\mathrm{x}_j(t) - \mathrm{y}_i(t)\| \leq d_{ij}$, where d_{ij} is a given positive real scalar.*

Remark 8. *It is worth noting that the scalar d_{ij} in Definition 9 is introduced to account for the physical size of target i and vehicle j in a practical setting, because every target and vehicle is mathematically describes as a point mass. For instance, if target 1 and vehicle 2 both have a spherical shape in $\mathbb{R}^d$ with radii r_1 and s_2, respectively, then $d_{12} = r_1 + s_2$.*

Definition 10. *For any $i \in \mathcal{I}_{\mathcal{T}}$, define the* first visit time *of target i, denoted by $\hat{\tau}_i \in \bar{\mathbb{R}}_{\geq 0}$,* *as the time the target is visited by one of the vehicles for the first time, i.e.*

$$\hat{\tau}_i = \inf\{t \in \mathbb{R}_{\geq 0} \ ; \ \min\{\|\mathrm{x}_j(t) - \mathrm{y}_i(t)\| - d_{ij}; j \in \mathcal{I}_{\mathcal{V}}\} \leq 0\}. \tag{4.7}$$

Note that for any $i \in \mathcal{I}_{\mathcal{T}}$, if the first visit time of target i is infinity, i.e. $\hat{\tau}_i = \infty$, this means that none of the vehicles visits that target.

Accordingly, one can define the set of indices of targets visited up to time t as follows

$$\hat{\mathcal{I}}_{\mathcal{T}}(t) = \{i \in \mathcal{I}_{\mathcal{T}} \mid \hat{\tau}_i \leq t\}. \tag{4.8}$$

Similarly, denote by $\mathcal{I}_{\mathcal{T}}(t)$ the set of targets arrived but not visited up to time t, i.e.

$$\mathcal{I}_{\mathcal{T}}(t) = \check{\mathcal{I}}_{\mathcal{T}}(t) \backslash \hat{\mathcal{I}}_{\mathcal{T}}(t) = \{i \in \mathcal{I}_{\mathcal{T}} \ ; \ \check{\tau}_i \leq t < \hat{\tau}_i\}. \tag{4.9}$$

For any $i \in \mathcal{I}_{\mathcal{T}}$, the trajectory of target i is assumed to be a function $\mathrm{y}_i : [\check{\tau}_i, \hat{\tau}_i] \to \mathbb{R}^d$, satisfying the local and global geometric conditions introduced next.

Assumption 6. *(Geometric Conditions)*

1. *(*Global Geometric Condition*) For any $i \in \mathcal{I}_{\mathcal{T}}$ and $\tau \in [\check{\tau}_i, \hat{\tau}_i]$, one has $\mathrm{y}_i(\tau) \in \mathcal{M}$, i.e. $\mathrm{y}_i([\check{\tau}_i, \hat{\tau}_i]) \subset \mathcal{M}$.*

2. *(*Local Geometric Condition*) For any $i \in \mathcal{I}_{\mathcal{T}}$, $\mathrm{y}_i : [\check{\tau}_i, \hat{\tau}_i] \to \mathcal{M}$ is a continuously differentiable function. Moreover, there exist non-negative scalars $\tilde{v}$ and $\tilde{a}$ such*

that for any $\tau \in [\check{\tau}_i, \hat{\tau}_i]$*, the following inequality holds*

$$\left\| \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{y}_i(\tau) \right\| \leq \tilde{v}. \tag{4.10}$$

Let $\alpha_i(t, \tau)$ *be a function satisfying the following equality*

$$\mathrm{y}_i(t) = \mathrm{y}_i(\tau) + \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{y}_i(\tau)(t - \tau) + \frac{1}{2} \alpha_i(t, \tau)(t - \tau)^2, \tag{4.11}$$

where $t \in [\tau, \hat{\tau}_i]$*. Then*

$$\sup_{s \in (\tau, \hat{\tau}_i]} \|\alpha_i(s, \tau)\| \leq \tilde{a}. \tag{4.12}$$

The global geometric condition ensures that once a target enters in the mission space, it will always remain in it. Note that not only does this property depend on the target trajectories, it also depends on the geometry of the mission space. The local geometric condition, on the other hand, ensures that the speed and acceleration of a target cannot exceed some prescribed values.

Assumption 7. *The position and velocity vectors of every target are available at the beginning of each time horizon, i.e.* τ *in (4.11).*

Assumption 7 provides grounds to estimate the position of target i at any future time $t \in [\tau, \hat{\tau}_i]$. For example, a first-order estimate is expressed as

$$\hat{\mathrm{y}}_i(t) = \mathrm{y}_i(\tau) + (t - \tau)\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(\tau). \tag{4.13}$$

Remark 9. *It follows immediately from* (4.12) *that*

$$\|\hat{\mathrm{y}}_i(t) - \mathrm{y}_i(t)\| \leq \frac{1}{2!}(t-\tau)^2 \sup_{t\in(\tau,\hat{\tau}_i]} \|\alpha_i(t,\tau)\| = \frac{1}{2!}(t-\tau)^2 \tilde{a} \tag{4.14}$$

for any $t \in [\tau, \hat{\tau}_i]$. *Thus, the closer* t *is to* τ, *the more precise the above estimation is.*

Corresponding to each target, one can define a *task* which is accomplished if one of the vehicles visits that target in finite time. By a harmless abuse of notation, let $\mathcal{I}_\mathcal{T}$ be the set of all tasks, $\check{\mathcal{I}}_\mathcal{T}(t)$ be the set of tasks started by time t, $\hat{\mathcal{I}}_\mathcal{T}(t)$ be the set of tasks accomplished by time t, and $\mathcal{I}_\mathcal{T}(t)$ be the set of tasks in process at time t. The *mission* is said to be accomplished when all tasks are accomplished.

Given the limitations of the vehicles in terms of information exchange and the unpredictable nature of the environment discussed earlier, it is desired to design a near-optimal cooperative control law in order to accomplish the mission. This problem will be investigated in the subsequent sections.

4.3 A Game-Theoretic Cooperative Receding Horizon Scheme

As an incentive for the vehicles to visit the targets, corresponding to each task a decreasing reward function is defined for every target which can be collected only if the task is accomplished (i.e., the target is visited). The vehicles dynamically make their decisions toward maximizing the total collected rewards. The decision-making process of every vehicle, which is iterative, consists of planning their paths and deciding upon their strategies for visiting the targets in an efficient fashion. At the beginning of each iteration, every vehicle updates its information by checking its sensing region, communicating with its neighbours, and then calculating the heading accordingly.

4.3.1 Reward Allocations

For any $i \in \mathcal{I}_{\mathcal{T}}$, let $\mathscr{R}_i$ be the *initial reward* considered for accomplishing task i. Define $\mathfrak{d}_i(\cdot) : \mathbb{R}_{\geq 0} \to [0, 1]$ as the *time discount function*, which is a decreasing function reflecting the rate of reward loss over time. The *reward function* of task i is then equal to $\mathscr{R}_i \mathfrak{d}_i(t)$, for any $i \in \mathcal{I}_{\mathcal{T}}$. There are various choices for the time discount function in order to model different aspects of timing and scheduling such as deadlines and priorities. As a simple example, one can consider the following function

$$\mathfrak{d}_i(t) = e^{-\gamma_i t}, \quad \forall i \in \mathcal{I}_{\mathcal{T}} \tag{4.15}$$

where $\gamma_i \in \mathbb{R}_{>0}$ is a parameter reflecting the degree of importance of target i.

4.3.2 Cooperative Structure

For any $i \in \mathcal{I}_{\mathcal{T}}$ and $j \in \mathcal{I}_{\mathcal{V}}$, the *assignment* of target i to vehicle j is characterized by a real scalar in $[0, 1]$, denoted by a_{ij}. This parameter reflects the level of interest of vehicle j in target i being assigned to it, and depends on the potential rewards to be collected as well as the positions of the vehicles and targets. At each step of the decision-making process, every vehicle is required to plan its task assignment based on available information. Note that these assignments highly depend on the sensing and communication capabilities of vehicles. More precisely, some of the vehicles may sense only a subset of targets, not all of them. Moreover, some of the vehicles may not be able to communicate with each other directly. These constraints need to be addressed in the assignment functions. To this end, the notion of *virtual targets* is introduced, and the definitions of target set and vehicle set are modified accordingly to take the communication and sensing limitations into account.

For any $j \in \mathcal{I}_{\mathcal{V}}$, denote by $\emptyset_j$ a *virtual target* which, if existed, it could be detected only by vehicle j. Let $\tilde{\mathcal{I}}_{\mathcal{T}}$ represent the set of all virtual targets, i.e. $\tilde{\mathcal{I}}_{\mathcal{T}} = \{\emptyset_j\}_{j\in\mathcal{I}_{\mathcal{V}}}$. Then, at any time $t \geq 0$, the set of targets in the sensing range of vehicle j, denoted by $\mathcal{I}_{\mathcal{T},j}(t)$, and the set of vehicles communicating with vehicle j, denoted by $\tilde{\mathcal{I}}_{\mathcal{V},j}(t)$, are defined respectively as

$$\mathcal{I}_{\mathcal{T},j}(t) = \{i \in \mathcal{I}_{\mathcal{T}}(t)\ ;\ \mathrm{y}_i(t) \in \mathscr{S}_j(t)\} \cup \{\emptyset_j\}, \tag{4.16}$$

and

$$\tilde{\mathcal{I}}_{\mathcal{V},j}(t) = \{\tilde{j} \in \mathcal{I}_{\mathcal{V}}\ ;\ \mathrm{x}_j(t) \in \mathscr{C}_{\tilde{j}}(t)\}. \tag{4.17}$$

Similarly, for any $i \in \mathcal{I}_{\mathcal{T}} \cup \tilde{\mathcal{I}}_{\mathcal{T}}$, the *set of sensing vehicles* for target i is defined below

$$\mathcal{I}_{\mathcal{V},i}(t) = \{j \in \mathcal{I}_{\mathcal{V}}\ ;\ i \in \mathcal{I}_{\mathcal{T},j}(t)\}. \tag{4.18}$$

One can similarly define the set of *sensible targets* as a group of targets, each of which lies in the sensing region of at least one of the vehicles. This set can be expressed as

$$\underline{\mathcal{I}}_{\mathcal{T}}(t) = \bigcup_{j\in\mathcal{I}_{\mathcal{V}}} \mathcal{I}_{\mathcal{T},j}(t). \tag{4.19}$$

Remark 10. *Since $\emptyset_j \in \mathcal{I}_{\mathcal{T},j}(t)$, for any $j \in \mathcal{I}_{\mathcal{V}}$, one can write*

$$\{\emptyset_j\ ;\ j \in \mathcal{I}_{\mathcal{V}}\} \subseteq \bigcup_{j\in\mathcal{I}_{\mathcal{V}}} \mathcal{I}_{\mathcal{T},j}(t) = \underline{\mathcal{I}}_{\mathcal{T}}(t), \tag{4.20}$$

which implies that $|\mathcal{I}_{\mathcal{V}}| \leq |\underline{\mathcal{I}}_{\mathcal{T}}(t)|$.

Note that at any point in time, each assignment depends on the positions of all vehicles and targets, for any $i \in \mathcal{I}_{\mathcal{T},j}(t)$ and $j \in \mathcal{I}_{\mathcal{V}}$. Thus, assignment a_{ij} can be expressed as a function of the following form:

$$a_{ij} : \mathcal{M}^{|\mathcal{I}_{\mathcal{V},i}(t)|} \times \mathcal{M}^{|\mathcal{I}_{\mathcal{T},j}(t)|} \rightarrow [0, 1]. \tag{4.21}$$

Note also that assignment a_{ij} depends on the vehicles that sense target i as well as the targets which are sensed by vehicle j, at any time $t \geq 0$.

The desired assignment is required to satisfy certain conditions. For example, each vehicle should normally consider all the targets inside its sensing region. Hence, for each vehicle, the sum of target assignments in its sensing region at any time t, should be equal to one, i.e.

$$\sum_{i \in \mathcal{I}_{\mathcal{T},j}(t)} a_{ij}(\mathbf{x}_i, \mathbf{y}_j) = 1, \quad \forall j \in \mathcal{I}_{\mathcal{V}}, \tag{4.22}$$

where $\mathbf{x}_i = (\mathrm{x}_{\tilde{j}})_{\tilde{j} \in \mathcal{I}_{\mathcal{V},i}(t)}$, $\mathbf{y}_j = (\mathrm{y}_i)_{i \in \mathcal{I}_{\mathcal{T},j}(t)}$ are the vectors of vehicles' positions and targets' positions, respectively. On the other hand, it is desired to accomplish as many tasks as possible, and also the number of targets in the sensing regions at any point in time is more than or equal to the number of vehicles. Therefore, with respect to to each vehicle, it is important to be conservative and under-assign the targets in its sensing regions to it, i.e.,

$$\sum_{j \in \mathcal{I}_{\mathcal{V},i}} a_{ij}(\mathbf{x}_i, \mathbf{y}_j) \leq 1, \quad \forall i \in \mathcal{I}_{\mathcal{T}} \cup \tilde{\mathcal{I}}_{\mathcal{T}}. \tag{4.23}$$

Definition 11. *Define* sensing bigraph*, denoted by $\mathcal{G}_t = (\mathcal{T}_t \cup \mathcal{V}_t, \mathcal{E}_t)$, as a bipartite graph with vertex partitions $\mathcal{T}_t = \underline{\mathcal{I}}_{\mathcal{T}}(t)$ and $\mathcal{V}_t = \mathcal{I}_{\mathcal{V}}$, and the edge set defined as $\mathcal{E}_t = \{(i,j) \in \mathcal{T}_t \times \mathcal{V}_t \ ; \ i \in \mathcal{I}_{\mathcal{T},j}(t)\}$. Let $\mathbf{B}_t$ be the biadjacency matrix of $\mathcal{G}_t$.*

For any $m, n \in \mathbb{N}$, $m \leq n$, define the set $\mathbb{A}^{n\times m}$ as $\{\mathbf{A} \in [0,1]^{n\times m}; \mathbf{A}^{\mathsf{T}}\mathbf{1}_n = \mathbf{1}_m, \mathbf{A}\mathbf{1}_m \leq \mathbf{1}_n\}$. Given a sensing bigraph $\mathcal{G}_t$, note that equations (4.22) and (4.23) introduce a set of constraints that any desired assignment $\mathbf{A}(\mathbf{x},\mathbf{y}) = (a_{ij}(\mathbf{x},\mathbf{y}))_{i\in\mathcal{I}_{\mathcal{T}}(t), j\in\mathcal{I}_{\mathcal{V}}}$ should satisfy them. More precisely, $\mathbf{A}(\mathbf{x},\mathbf{y})$ belongs to the set $\mathscr{A}_{\mathcal{I}_{\mathcal{T}}(t),\mathcal{I}_{\mathcal{V}}}$ defined as

$$\mathscr{A}_{\mathcal{I}_{\mathcal{T}}(t),\mathcal{I}_{\mathcal{V}}} = \{\mathbf{A} : \mathcal{M}^{|\mathcal{I}_{\mathcal{V}}|} \times \mathcal{M}^{|\underline{\mathcal{I}}_{\mathcal{T}}(t)|} \to \mathbb{A}^{|\underline{\mathcal{I}}_{\mathcal{T}}(t)| \times |\mathcal{I}_{\mathcal{V}}|}, \mathbf{A} \leq \mathbf{B}_t\}.$$

4.3.3 Cooperative Receding Horizon Trajectory Construction

It is desired now to develop a cooperative receding horizon controller (CRHC), which iteratively generates a set of headings, step sizes and optimal assignments for each vehicle such that the final collected rewards are maximized. The controller is applied at time instants denoted by $\{t_k\}_{k=0}^{\infty} \in \mathbb{R}_{\geq 0}$, where an optimization problem, estimating the collectible rewards in the future, is solved at each time instant. The solution of the optimization problem is based on currently available information, which are the current positions of the targets and vehicles, along with the predicted future positions of the targets. The solution of the optimization problem is used to obtain the optimal control input $\mathbf{u}^k = (\mathrm{u}_j(t_k))_{j\in\mathcal{I}_\mathcal{V}}$ as well as the optimal assignments $\{a_{ij}(\mathrm{x}(t_{k+1}), \hat{\mathrm{y}}(t_{k+1}))\}_{i\in\mathcal{I}_\mathcal{T}(t_k), j\in\mathcal{I}_\mathcal{V}}$.

Let the action horizon of CRHC be denoted by H_k (Note that H_k is a strictly positive real scalar). For any $j \in \mathcal{I}_\mathcal{V}$, apply the control input $\mathrm{u}_j(t_k)$ to vehicle j, in the time interval $(t_k, t_k + H_k)$. Then, it follows from equation (4.3) that at time $t_k + H_k$, the position of vehicle j is

$$\mathrm{x}_j(t_k + H_k) = \mathrm{x}_j(t_k) + \mathrm{u}_j(t_k)H_k, \tag{4.24}$$

for any $j \in \mathcal{I}_\mathcal{V}$. Similarly, based on the available information at time instant t_k, one can use equation (4.13) to estimate the position of target i at time instant $t_k + H_k$ as

$$\hat{\mathrm{y}}_i(t_k + H_k) = \mathrm{y}_i(t_k) + H_k \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(t_k), i \in \mathcal{I}_\mathcal{T}(t_k). \tag{4.25}$$

Denote by $\mathcal{T}_{\min,k}$ the earliest time that the next target can be visited, using the estimates obtained based on the information available at time t_k, i.e.

$$\mathcal{T}_{\min,k} = t_k + \min_{i\in\mathcal{I}_\mathcal{T}(t_k), j\in\mathcal{I}_\mathcal{V}} \|\mathrm{x}_j(t_k) - \mathrm{y}_i(t_k)\|(u_{\max} + \tilde{v})^{-1}.$$

The CRHC path planning continues until either the next immediate target is visited or a new target arrives, and then updates the position information of the targets. Thus, the following relation holds

$$H_k \leq \eta_k(\underline{\tau}_{\min,k} - t_k) = \eta_k \min_{i \in \mathcal{I}_{\mathcal{T}}(t_k), j \in \mathcal{I}_{\mathcal{V}}} \frac{\|\mathrm{x}_j(t_k) - \mathrm{y}_i(t_k)\|}{u_{\max} + \tilde{v}}, \tag{4.26}$$

where $\eta_k \in (0,1)$ is a coefficient used to model uncertainties. For simplicity, let the action horizon be chosen equal to the planning horizon. Therefore, $t_{k+1} = t_k + H_k$, which by substituting in equation (4.26) yields

$$\underline{\tau}_{\min,k} - t_{k+1} = (\underline{\tau}_{\min,k} - t_k) - H_k > 0. \tag{4.27}$$

The position of target i at any time $t \geq \underline{\tau}_{\min,k}$ is

$$\mathrm{y}_i(t) = \mathrm{y}_i(t_k) + \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(t_k)(t - t_k) + \frac{1}{2}\alpha_i(t, t_k)(t - t_k)^2, \tag{4.28}$$

for any $i \in \mathcal{I}_{\mathcal{T}}(t_k)$. Now, it follows from equation (4.25) that

$$\mathrm{y}_i(t) = \hat{\mathrm{y}}_i(t_{k+1}) + \frac{1}{2}\mathscr{E}_t^k (t - t_k)^2 \tilde{\boldsymbol{a}}, \tag{4.29}$$

where $\mathscr{E}_t^k$ is defined as

$$\mathscr{E}_t^k = \alpha_i(t, t_k)\tilde{\boldsymbol{a}}^{-1} + 2(t - t_{k+1})\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(t_k)[(t - t_k)^2 \tilde{\boldsymbol{a}}]^{-1}.$$

Since $t - t_k > t - t_{k+1}$, it is concluded that

$$\|\mathscr{E}_t^k\| < 1 + 2\|\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(t_k)\|[(t - t_k)\tilde{\boldsymbol{a}}]^{-1}. \tag{4.30}$$

The second term in the right hand side of equation (4.30) satisfies the following inequality

$$\|\frac{2\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(t_k)}{(t - t_{k+1})\tilde{\boldsymbol{a}}}\| \leq \|\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(t_k)\|\frac{2}{(\mathcal{I}_{\min,k} - t_k)\tilde{\boldsymbol{a}}}. \tag{4.31}$$

Note that the denominator in the right side of (4.31) is sufficiently large if the targets and vehicles are very far from each other, or $\tilde{\boldsymbol{a}}$ is sufficiently large. In that case, the right hand side of (4.31) will be negligible, and hence

$$\|2\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(t_k)[(t - t_{k+1})\tilde{\boldsymbol{a}}]^{-1}\| \ll 1. \tag{4.32}$$

Note that while the trajectory of target i is *a priori* unknown, it is uniformly distributed and bounded by $\tilde{\boldsymbol{a}}$. Define $\bar{\mathscr{E}}_t^k = \tilde{\boldsymbol{a}}^{-1}\alpha_i(t, t_k)$, and let it be a uniformly distributed random vector, taking magnitudes between 0 and $\tilde{\boldsymbol{a}}$ and different directions, such that (4.30) and (4.32) yield $\mathbb{E}[\mathscr{E}_t^k] = \mathbf{0}_{|\mathcal{I}_\mathcal{T}(t)|}$. Using this equality and (4.29), one can estimate y_i for large values of t as $\hat{\mathrm{y}}_i(t_{k+1})$. From this estimation and also the current positions of targets and vehicles as well as the control input u^k, the time that vehicle j visits target i can be estimated as

$$\hat{\tau}_{ij}(\mathrm{u}^k, t_k) = (t_k + H_k) + \|\mathrm{x}_j(t_{k+1}) - \hat{\mathrm{y}}_i(t_{k+1})\| u_{\max}^{-1}, \tag{4.33}$$

for any $j \in \mathcal{I}_{\mathcal{V}}$. Note that in the above equation it is assumed that the control input u_j remains unchanged after the time instant t_{k+1} until the vehicle reaches the position of target $\hat{\mathrm{y}}_i(t_{k+1})$. Note also that CRHC updates the estimates (including the ones given above) in each iteration to reduce the estimation error. Similarly, if $a_{ij}(\mathbf{x}(t_{k+1}), \hat{\mathbf{y}}(t_{k+1}))$ is the optimal assignment, regardless of uncertainties, one can expect that this assignment remains unchanged until vehicle j visits target i. Therefore

$$a_{ij}(\mathbf{x}(\hat{\tau}_{ij}(\mathbf{u}^k, t_k)), \hat{\mathbf{y}}(\hat{\tau}_{ij}(\mathbf{u}^k, t_k))) = a_{ij}(\mathbf{x}(t_{k+1}), \hat{\mathbf{y}}(t_{k+1})).$$

Consequently, one can estimate the maximum total reward which the vehicles are expected at time t_{k+1} to collect by the end of the mission. Denote by $\mathfrak{R}^{k+1}$ this estimated expected reward. For simplicity of notations, let $\tilde{\mathfrak{d}}_{ij}(\mathbf{u}^k, t_k) = \mathfrak{d}_i[\hat{\tau}_{ij}(\mathbf{u}^k, t_k)]$ and $\tilde{a}_{ij}(\mathbf{u}^k, t_k) = a_{ij}(\mathbf{x}(\hat{\tau}_{ij}(\mathbf{u}^k, t_k)), \hat{\mathbf{y}}(\hat{\tau}_{ij}(\mathbf{u}^k, t_k)))$. Then

$$\mathfrak{R}^{k+1}(\mathbf{u}^k, t_k) = \sum_{j \in \mathcal{I}_{\mathcal{V}}} \sum_{i \in \mathcal{I}_{\mathcal{T},j}(t_k)} \mathcal{R}_i \tilde{\mathfrak{d}}_{ij}(\mathbf{u}^k, t_k) \tilde{a}_{ij}(\mathbf{u}^k, t_k). \tag{4.34}$$

From (4.43), the k^{th} iteration in the CRHC, represented by P_k, can be written as

$$\mathrm{P}_k: \quad \begin{cases} \max & \mathfrak{R}^{k+1}(\mathbf{u}^k, t_k) \\ \text{s.t.} & \tilde{\mathbf{A}}(\mathbf{u}^k, t_k) \in \mathscr{A}^k, \mathbf{u}^k \in \mathcal{U}^k, \end{cases} \tag{4.35}$$

where $\mathscr{A}^k = \mathscr{A}_{\mathcal{I}_\mathcal{T}(t_k),\mathcal{I}_\mathcal{V}(t_k)}$ and $\mathcal{U}^k = \{\mathbf{u} = (\mathrm{u}_j)_{j\in\mathcal{I}_\mathcal{V}} \ ; \ \mathrm{u}_j \in \mathbb{R}^d, \|\mathrm{u}_j\| \leq u_{\max}, \forall j \in \mathcal{I}_\mathcal{V}\}$ is the set of *admissible heading control*.

For convenience of notation, $\mathrm{x}_j(t_k)$, $\mathrm{y}_i(t_k)$ and $\hat{\mathrm{y}}_i(t_k)$ will hereafter be denoted by x_j^k, y_i^k and $\hat{\mathrm{y}}_i^k$, respectively, for any $i \in \mathcal{I}_\mathcal{T}$ and $j \in \mathcal{I}_\mathcal{V}$.

4.3.4 Extension To Game Theoretic Formulation

It is desired now to develop a distributed cooperative receding horizon controller (DCRHC) based on the proposed CRHC. For simplicity, assume that there is no target priority and set for any $i \in \mathcal{I}_\mathcal{T}$, $\gamma_i = \gamma$ and $\mathcal{R}_i = \mathcal{R}$, for some $\gamma, \mathcal{R} \in \mathbb{R}_{\geq 0}$.

Theorem 6. *Consider the performance index* $J_{\hat{\mathbf{y}}^{k+1}_{\bullet\mathcal{I}_\mathcal{T}(t)}} : \mathcal{M}^{|\mathcal{I}_\mathcal{V}|} \times \mathbb{A}^{|\mathcal{I}_\mathcal{T}(t)|\times|\mathcal{I}_\mathcal{V}|} \to \mathbb{R}_{\geq 0}$ *as*

$$J_{\hat{\mathbf{y}}^{k+1}_{\bullet\mathcal{I}_\mathcal{T}(t)}}(\mathbf{x}, \mathbf{A}) = \sum_{j\in\mathcal{I}_\mathcal{V}} \sum_{i\in\mathcal{I}_{\mathcal{T},j}(t)\backslash\tilde{\mathcal{I}}_\mathcal{T}} a_{ij} e^{-\bar{\gamma}\|\mathrm{x}_j - \hat{\mathrm{y}}_i^{k+1}\|}, \tag{4.36}$$

where $\bar{\gamma} = \gamma u_{\max}^{-1}$. *Then, if* $\mathcal{I}_\mathcal{T}(t_k)\backslash\tilde{\mathcal{I}}_\mathcal{T} \neq \emptyset$, *the optimization problem* P^k *presented in*

(4.35) is equivalent to

$$\begin{aligned} \max \quad & J_{\hat{\mathbf{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}), \\ \text{s.t.} \quad & \mathbf{A} \in \mathbb{A}^{|\underline{\mathcal{I}}_{\mathcal{T}}(t_k)| \times |\mathcal{I}_{\mathcal{V}}|}, \\ & \|\mathrm{x}_j - \mathrm{x}_j^k\| \leq u_{\max} H_k, \forall j \in \mathcal{I}_{\mathcal{V}}. \end{aligned} \tag{4.37}$$

Proof. See Section 4.5.1. □

For any $j \in \mathcal{I}_{\mathcal{V}}$ and $i \in \underline{\mathcal{I}}_{\mathcal{T}}(t_k)$, denote by a_j the j^{th} column of $\mathbf{A}$ and by $\mathrm{A}_i^{\mathsf{T}}$ its i^{th} row. Define the *penalty function* $p(\mathrm{A}) = \max(0, \mathrm{A}^{\mathsf{T}} \mathbf{1}_{|\mathcal{I}_{\mathcal{V}}|} - 1)$, for any $\mathrm{A} \in \mathbb{R}^{|\mathcal{I}_{\mathcal{V}}|}$. One can show that as $\lambda \to \infty$, the solution of the following maximization problem

$$\begin{aligned} \max \quad & J_{\hat{\mathbf{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}) - \lambda \textstyle\sum_{i \in \underline{\mathcal{I}}_{\mathcal{T}}(t_k) \backslash \tilde{\mathcal{I}}_{\mathcal{T}}} p(\mathrm{A}_i), \\ \text{s.t.} \quad & \|\mathrm{x}_j - \mathrm{x}_j^k\| \leq u_{\max} H_k, \quad \forall j \in \mathcal{I}_{\mathcal{V}}, \\ & \mathrm{a}_j^{\mathsf{T}} \mathbf{1}_{|\underline{\mathcal{I}}_{\mathcal{T}}(t)|} = 1, \quad \forall j \in \mathcal{I}_{\mathcal{V}}, \\ & \mathrm{a}_j \geq \mathbf{0}_{|\underline{\mathcal{I}}_{\mathcal{T}}(t)|}, \quad \forall j \in \mathcal{I}_{\mathcal{V}}, \\ & \mathbf{A} \leq \mathbf{B}_{t_k}. \end{aligned} \tag{4.38}$$

converges to that of (4.37). Now, for any $j \in \mathcal{I}_{\mathcal{V}}$ and $i \in \mathcal{I}_{\mathcal{T}(t_k),j} \backslash \tilde{\mathcal{I}}_{\mathcal{T}}$, define $\hat{\mathrm{x}}_{j \to i}^{k+1}$ as

$$\hat{\mathrm{x}}_{j \to i}^{k+1} = \mathrm{x}_j^k + \frac{\hat{\mathrm{y}}_i^{k+1} - \mathrm{x}_j^k}{\|\hat{\mathrm{y}}_i^{k+1} - \mathrm{x}_j^k\|} u_{\max} H_k, \tag{4.39}$$

which is, in fact, the future position of vehicle j as it is aimed to move towards predicted position of target i, and let

$$d_{ij}^k = e^{-\bar{\gamma}\|\hat{\mathrm{x}}_{j\to i}^{k+1} - \hat{\mathrm{y}}_i^{k+1}\|}. \tag{4.40}$$

For any t_k, define the finite game $\mathfrak{G}_k = (\mathcal{I}_\mathcal{V}, \times_{j\in\mathcal{I}_\mathcal{V}} \mathcal{I}_{\mathcal{T}(t_k),j}, \mathbf{U} = (U_j)_{j\in\mathcal{I}_\mathcal{V}})$, where

$$U_j\big((\mathrm{a}_{\tilde{\jmath}})_{\tilde{\jmath}\in\mathcal{I}_\mathcal{V}}\big) = \sum_{i\in\mathcal{I}_{\mathcal{T}(t_k),j}\setminus\tilde{\mathcal{I}}_\mathcal{T}} \big(\sum_{\tilde{\jmath}\in\mathcal{I}_{\mathcal{V},i}} a_{i\tilde{\jmath}} d_{i\tilde{\jmath}}^k - \lambda p((a_{i\tilde{\jmath}})_{\tilde{\jmath}\in\mathcal{I}_\mathcal{V}})\big).$$

Note that the set of action profiles here is the same as the set of assignments with values 0 or 1. Similarly, it can be verified that each assignment is a strategy profile for the game $\mathfrak{G}_k$.

Theorem 7. *The game $\mathfrak{G}_k$ is a* potential game *with the following potential function*

$$P\big((\mathrm{a}_{\tilde{\jmath}})_{\tilde{\jmath}\in\mathcal{I}_\mathcal{V}}\big) = \sum_{j\in\mathcal{I}_\mathcal{V}} \sum_{i\in\mathcal{I}_{\mathcal{T},j}(t_k)\setminus\tilde{\mathcal{I}}_\mathcal{T}} a_{ij} e^{-\bar{\gamma}\|\hat{\mathrm{x}}_{j\to i}^{k+1} - \hat{\mathrm{y}}_i^{k+1}\|} - \lambda \sum_{i\in\underline{\mathcal{I}}_\mathcal{T}(t_k)} p((a_{i\tilde{\jmath}})_{\tilde{\jmath}\in\mathcal{I}_\mathcal{V}}). \tag{4.41}$$

Proof. See Section 4.5.2. □

Theorem 8. *There exists a constant $\underline{\lambda}$ such that for any $\lambda \geq \underline{\lambda}$, the optimization problem* (4.38) *has a solution* $(\mathbf{x}_\lambda^*, \mathbf{A}_\lambda^*)$ *with the entries of* $\mathbf{A}_\lambda^*$ *being* 0 *or* 1 *which is a solution of* (4.37)*. Furthermore, this solution,* $\mathbf{A}_\lambda^*$*, is a maximizer for the potential function* (4.41)*, and hence a pure Nash equilibrium for* $\mathfrak{G}_k$*.*

Proof. See Section 4.5.3. □

Remark 11. *Since the vehicles are capable of communicating with each other, they can share with their neighbors their actions on the targets located in the intersection of their sensing regions. Based on this information exchange, a method such as* generalized regret monitoring *(GRM) [132] or* spatial adaptive play *(SAP) [133] which guarantees sufficiently fast convergence to a pure Nash equilibrium can be applied to obtain a Nash equilibrium.*

4.4 Simulation Results

In this section, the performance of the proposed method with GRM and SAP dynamic learning approaches is investigated by simulations involving two vehicles and a set of five targets arriving sequentially in the mission space. The sensing range for both vehicles is $r_s = 5$m, and the mission space is a closed convex set in the plane $\mathcal{M} = [-20, 20] \times [-20, 20]\text{m}^2$. Targets have *a priori* unknown trajectories (randomly chosen in the simulation) with the maximum velocity $\tilde{v} = 1.5$m/s and the upper bound on the magnitude of vehicles' velocity is $u_{\max} = 2$m/s. Initially, along with the two vehicles, two targets are also present in the mission space, and the remaining three targets arrive sequentially at 3s, 4s, 6s.

Case 1 (SAP)**.** *In this learning method [133], vehicles negotiate with each other to reach the pure Nash equilibrium by computing a utility function, where intercepting a target is rewarded while selection of one target by more than one vehicle is penalized by*

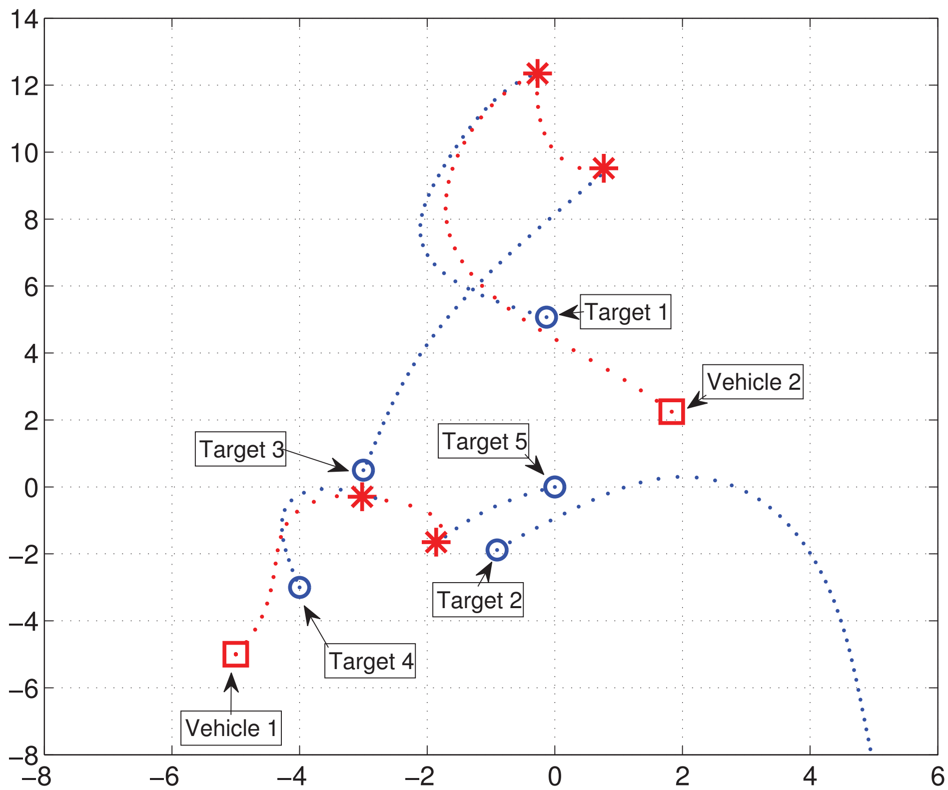

Figure 4.1: An example of target tracking with two vehicles and sequentially arriving targets, using the spatial adaptive play (SAP) as a game learning mechanism.

a negative term with a sufficiently large magnitude. Figure 4.1 shows the result of this learning mechanism, where it can be observed that other than target 2 *that has been out of the sensing region of the vehicles, all other targets are intercepted by a vehicle in a cooperative manner.*

Case 2 (GRM)**.** *In the GRM learning method [132], a fading memory and inertia mechanism are also utilized to enable fast convergence to pure Nash equilibrium, where the*

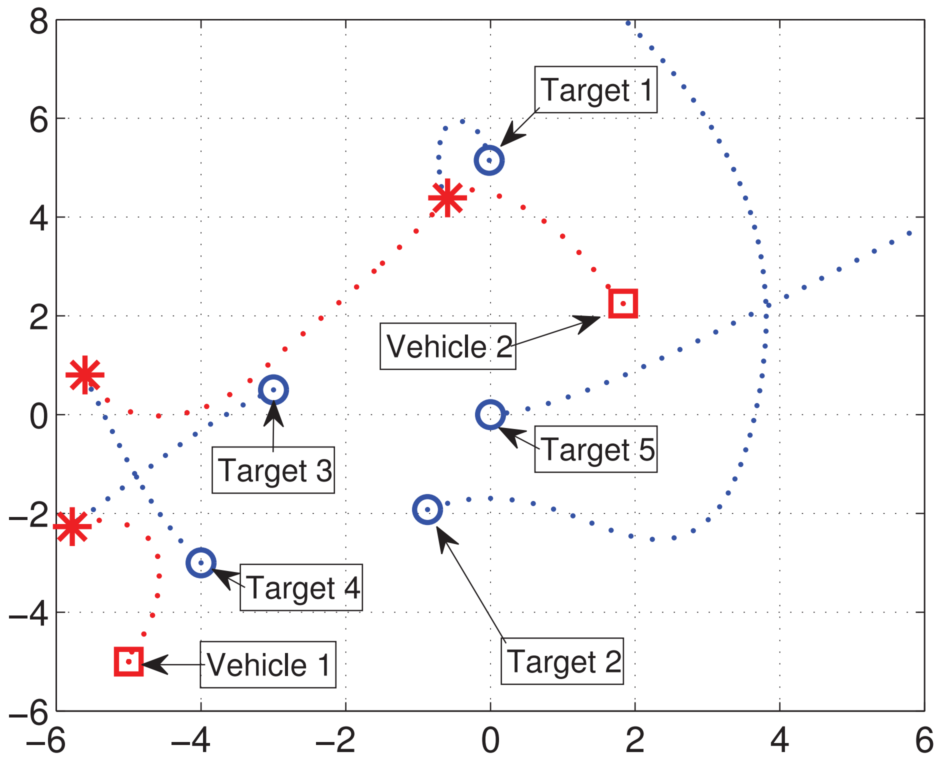

Figure 4.2: An example of target tracking with two vehicles and a set of five sequentially arriving targets, using the generalized regret monitoring (GRM).

forgetting factor is set to $\rho = 0.99$ *and the inertia is* $\alpha = 0.95$. *Figure 4.2 depicts the result of this simulation, which demonstrates that targets* 2 *and* 5 *do not enter the sensing region of any of the two vehicles but all other targets are intercepted by the vehicles. As a result of negotiation in the game played by two vehicles in this example, they switch their selected targets* 3 *and* 4 *at some point in time. It is worth noting that the vehicles stop moving when there is no target in their sensing region. This is a result of the extra virtual target that was added for each vehicle which can be selected when no target is sensible.*

4.5 Appendices

4.5.1 Proof of Theorem 6

For any fixed $H_k \in \mathbb{R}_{\geq 0}$, the equation (4.24) gives a one-to-one correspondence between $\mathcal{U}^k$ and $\mathscr{B}(\mathbf{x}^k, u_{\max} H_k)$. Considering (4.15) and (4.33), for any $i \in \underline{\mathcal{I}}_{\mathcal{T}}(t)$ and any $j \in \mathcal{I}_{\mathcal{V}}$, one has

$$\begin{aligned} \tilde{\mathfrak{d}}_{ij}(\mathbf{u}^k, t_k) &= e^{-\gamma_i (t_k + H_k + \frac{\|\mathrm{x}_j^{k+1} - \hat{\mathrm{y}}_i^{k+1}\|}{u_{\max}})} \\ &= e^{-\gamma_i (t_k + H_k)} e^{-\gamma_i u_{\max}^{-1} \|\mathrm{x}_i^{k+1} - \hat{\mathrm{y}}_i^{k+1}\|}. \end{aligned} \tag{4.42}$$

Since for any $i \in \mathcal{I}_{\mathcal{T}}$, it is assumed that $\gamma_i = \gamma$ and $\mathcal{R}_i = \mathcal{R}$, it yields

$$\mathfrak{R}^{k+1} = \mathcal{R} e^{-\gamma(t_k + H_k)} \sum_{j \in \mathcal{I}_{\mathcal{V}}} \sum_{i \in \mathcal{I}_{\mathcal{T},j}(t_k)} \tilde{a}_{ij} e^{-\bar{\gamma} \|\mathrm{x}_i^{k+1} - \hat{\mathrm{y}}_i^{k+1}\|}, \tag{4.43}$$

where the arguments $(\mathbf{u}^k, t_k)$ are omitted for brevity. Hence, from (4.36), it concludes that

$$\mathfrak{R}^{k+1}(\mathbf{u}^k, t_k) = \mathcal{R} e^{-\gamma(t_k + H_k)} J_{\hat{\mathbf{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t)}}(\mathbf{x}^{k+1}, \mathbf{A}^k). \tag{4.44}$$

Therefore, as $\mathcal{R} e^{-\gamma(t_k + H_k)} > 0$, the optimization problem (4.35) is equivalent with

$$\begin{aligned} \max \quad & J_{\hat{\mathbf{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^{k+1}, \mathbf{A}^k), \\ \text{s.t.} \quad & \mathbf{A}^k \in \mathbb{A}^{|\underline{\mathcal{I}}_{\mathcal{T}}(t_k)| \times |\mathcal{I}_{\mathcal{V}}|}, \\ & \|\mathrm{x}_j^{k+1} - \mathrm{x}_j^k\| \leq u_{\max} H_k, \forall j \in \mathcal{I}_{\mathcal{V}}. \end{aligned} \tag{4.45}$$

Changing names of the variables, the optimization problem (4.37) yields.

4.5.2 Proof of Theorem 7

Consider vehicle j. Let i', i'' be indices of two targets in $\mathcal{I}_{\mathcal{T}(t_k),j}$ and the standard vectors $\mathrm{e}_{i'}, \mathrm{e}_{i''} \in \mathbb{R}^{|\mathcal{I}_{\mathcal{T}(t_k)}|}$ be the respective action vectors, i.e. $\mathrm{a}'_j = \mathrm{e}_{i'}$ and $\mathrm{a}''_j = \mathrm{e}_{i''}$. Also, let $\mathrm{a}_{-j} \in \times_{j \neq i} \mathbb{R}^{|\underline{\mathcal{I}}_{\mathcal{T}}(t_k)|}$ represents actions for the vehicles with indices in $\mathcal{I}_{\mathcal{V}} \backslash \{j\}$, i.e. $\mathrm{a}_{-j} = (\mathrm{a}_l)_{l \in \mathcal{I}_{\mathcal{V}} \backslash \{j\}}$ where $\mathrm{a}_l = \mathrm{e}_{i_l}$ is the action vector for vehicle l and $i_l \in \mathcal{I}_{\mathcal{T}(t_k),l}$ is the target with respect to action vector a_l, for any $l \in \mathcal{I}_{\mathcal{V}} \backslash \{j\}$. Now, set $\mathbf{A}' = (a'_{ij})$ as $(\mathrm{a}'_j, \mathrm{a}_{-j})$ and $\mathbf{A}'' = (a''_{ij})$ as $(\mathrm{a}''_j, \mathrm{a}_{-j})$. In order to show that $\mathfrak{G}_k$ is a potential game, one needs to verify that

$$P(\mathrm{a}'_j, \mathrm{a}_{-j}) - P(\mathrm{a}''_j, \mathrm{a}_{-j}) = U_j(\mathrm{a}'_j, \mathrm{a}_{-j}) - U_j(\mathrm{a}''_j, \mathrm{a}_{-j}).$$

First, let $i', i'' \notin \tilde{\mathcal{I}}_{\mathcal{T}}$. From (4.41) one has

$$P(\mathbf{A}') - P(\mathbf{A}'') = \Big(\sum_{l \in \mathcal{I}_{\mathcal{V}}} \sum_{i \in \mathcal{I}_{\mathcal{T},l}(t_k) \backslash \tilde{\mathcal{I}}_{\mathcal{T}}} a'_{il} d^k_{il} - \lambda \sum_{i \in \underline{\mathcal{I}}_{\mathcal{T}}(t_k)} p((a'_{il})_{l \in \mathcal{I}_{\mathcal{V}}}) \Big) - \Big(\sum_{l \in \mathcal{I}_{\mathcal{V}}} \sum_{i \in \mathcal{I}_{\mathcal{T},l}(t_k) \backslash \tilde{\mathcal{I}}_{\mathcal{T}}} a''_{il} d^k_{il} - \lambda \sum_{i \in \underline{\mathcal{I}}_{\mathcal{T}}(t_k)} p((a''_{il})_{l \in \mathcal{I}_{\mathcal{V}}}) \Big). \tag{4.46}$$

Since, $\mathbf{A}'$ and $\mathbf{A}''$ differ only in j^{th} column and in rows $i', i'' \in \mathcal{I}_{\mathcal{T}(t_k),j}$, it yields that

$$P(\mathbf{A}') - P(\mathbf{A}'') = d^k_{i'j} - d^k_{i''j} - \lambda\Big(p((a'_{i'l})_{l\in\mathcal{I}_\mathcal{V}}) - p((a''_{i''l})_{l\in\mathcal{I}_\mathcal{V}})\Big). \tag{4.47}$$

Similarly, from (4.3.4) one has

$$U_j(\mathbf{A}') - U_j(\mathbf{A}'') = \Big(\sum_{i\in\mathcal{I}_{\mathcal{T}(t_k),j}\setminus\tilde{\mathcal{I}}_\mathcal{T}} \sum_{l\in\mathcal{I}_{\mathcal{V},i}} a'_{il}d^k_{il} - \lambda\sum_{i\in\mathcal{I}_{\mathcal{T}(t_k),j}\setminus\tilde{\mathcal{I}}_\mathcal{T}} p((a'_{il})_{l\in\mathcal{I}_\mathcal{V}})\Big) - \Big(\sum_{i\in\mathcal{I}_{\mathcal{T}(t_k),j}\setminus\tilde{\mathcal{I}}_\mathcal{T}} \sum_{l\in\mathcal{I}_{\mathcal{V},i}} a''_{il}d^k_{il} - \lambda\sum_{i\in\mathcal{I}_{\mathcal{T}(t_k),j}\setminus\tilde{\mathcal{I}}_\mathcal{T}} p((a''_{il})_{l\in\mathcal{I}_\mathcal{V}})\Big), \tag{4.48}$$

and also, as $\mathbf{A}'$ and $\mathbf{A}''$ differ only in j^{th} column and in rows $i', i'' \in \mathcal{I}_{\mathcal{T}(t_k),j}$, one can see that

$$U_j(\mathbf{A}') - U_j(\mathbf{A}'') = d^k_{i'j} - d^k_{i''j} - \lambda\Big(p((a'_{i'l})_{l\in\mathcal{I}_\mathcal{V}}) - p((a''_{i''l})_{l\in\mathcal{I}_\mathcal{V}})\Big). \tag{4.49}$$

From (4.47) and (4.49), it concludes that

$$P(\mathrm{a}'_j, \mathrm{a}_{-j}) - P(\mathrm{a}''_j, \mathrm{a}_{-j}) = U_j(\mathrm{a}'_j, \mathrm{a}_{-j}) - U_j(\mathrm{a}''_j, \mathrm{a}_{-j}). \tag{4.50}$$

In the case that $i' \in \tilde{\mathcal{I}}_\mathcal{T}$ or $i'' \in \tilde{\mathcal{I}}_\mathcal{T}$, with a similar discussion one can show that (4.50) holds. This proves that the game $\mathfrak{G}_k$ is a potential game.

4.5.3 Proof of Theorem 8

Preliminary Definitions and Theorems

Let $n, m \in \mathbb{N}$ and $\mathcal{G} = (\mathcal{U} \cup \mathcal{V}, \mathcal{E})$ be a bipartite graph with vertex partitions $\mathcal{U}$ and $\mathcal{V}$ where $|\mathcal{U}| = n$ and $|\mathcal{V}| = m$. Also, let $\mathbf{B}_{\mathcal{G}}$ be biadjacency matrix of the bipartite graph $\mathcal{G}$. Define

$$\mathbb{A}_{\mathcal{G}} = \{\mathbf{A} \in [0,1]^{n\times m}; \mathbf{A} \in \mathbb{A}^{n\times m}, \mathbf{A} \leq \mathbf{B}_{\mathcal{G}}\}, \tag{4.51}$$

and

$$\mathbb{B}_{\mathcal{G}} = \mathbb{A}_{\mathcal{G}} \cap \{0,1\}^{n\times m}. \tag{4.52}$$

Similarly, define

$$\tilde{\mathbb{A}}_{\mathcal{G}} = \{\mathbf{A} \in [0,1]^{n\times m}; \mathbf{A} \leq \mathbf{B}_{\mathcal{G}}, \mathbf{A}^{\mathsf{T}}\mathbf{1}_n = \mathbf{1}_m\}, \tag{4.53}$$

and

$$\tilde{\mathbb{B}}_{\mathcal{G}} = \tilde{\mathbb{A}}_{\mathcal{G}} \cap \{0,1\}^{n\times m}. \tag{4.54}$$

Now, let $\mathcal{I}$ be a set of natural numbers such that $\mathcal{I} \subseteq \mathbb{N}_n$ and define

$$\begin{aligned}\mathbb{A}_{\mathcal{G},\mathcal{I}} = \{\mathbf{A} \in [0,1]^{n\times m};\ & \mathbf{A} \leq \mathbf{B}_{\mathcal{G}}, \mathbf{A}^{\mathsf{T}}\mathbf{1}_n = \mathbf{1}_m, \\ & \mathbf{A} = (\mathrm{A}_i)^{\mathsf{T}}_{i\in\mathbb{N}_n}, \\ & \forall i \in \mathcal{I}, \mathrm{A}_i^{\mathsf{T}}\mathbf{1}_m \geq 1, \\ & \forall i \in \mathbb{N}_n \backslash \mathcal{I}, \mathrm{A}_i^{\mathsf{T}}\mathbf{1}_m \leq 1\},\end{aligned} \tag{4.55}$$

and

$$\mathbb{B}_{\mathcal{G},\mathcal{I}} = \mathbb{A}_{\mathcal{G},\mathcal{I}} \cap \{0,1\}^{n\times m}. \tag{4.56}$$

Theorem 9. *For any bipartite graph $\mathcal{G}$, one has that*

i. $\mathbb{A}_{\mathcal{G}} = \operatorname{conv} \mathbb{B}_{\mathcal{G}}$,

ii. $\tilde{\mathbb{A}}_{\mathcal{G}} = \operatorname{conv} \tilde{\mathbb{B}}_{\mathcal{G}}$,

iii. $\mathbb{A}_{\mathcal{G},\mathcal{I}} = \operatorname{conv} \mathbb{B}_{\mathcal{G},\mathcal{I}}$.

Proof of Theorem 8

Define function $\tilde{J}_{\hat{\mathbf{y}}^{k+1}_{\bullet\mathcal{I}_{\mathcal{T}}(t_k)}} : \mathcal{M}^{|\mathcal{I}_{\mathcal{V}}|} \times [0,1]^{|\mathcal{I}_{\mathcal{T}}(t_k)|\times|\mathcal{I}_{\mathcal{V}}|} \to \mathbb{R}_{\geq 0}$ as

$$\tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet\mathcal{I}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}) = J_{\hat{\mathbf{y}}^{k+1}_{\bullet\mathcal{I}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}) - \lambda \sum_{i\in\mathcal{I}_{\mathcal{T}}(t_k)\backslash\tilde{\mathcal{I}}_{\mathcal{T}}} p(\mathrm{A}_i), \tag{4.57}$$

and note that

$$J_{\hat{\mathbf{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}) \;=\; \sum_{i \in \underline{\mathcal{I}}_{\mathcal{T}}(t_k) \backslash \tilde{\mathcal{I}}_{\mathcal{T}}} \sum_{j \in \mathcal{I}_{\mathcal{V},i}(t_k)} a_{ij} e^{-\bar{\gamma}\|\mathrm{x}_j - \hat{\mathrm{y}}_i^{k+1}\|}. \tag{4.58}$$

Denote by Ω_{t_k} as $\times_{j=1}^{|\mathcal{I}_{\mathcal{V}}|} \mathscr{B}(\mathrm{x}_j^k, H_k u_{\max})$. Take $(\mathbf{x}^*, \mathbf{A}^*)$ such that

$$(\mathbf{x}^*, \mathbf{A}^*) \;\in\; \underset{(\mathbf{x},\mathbf{A}) \in \Omega_{t_k} \times \tilde{\mathbb{A}}_{\mathcal{G}_{t_k}}}{\operatorname{argmax}} \; \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}), \tag{4.59}$$

and let $\mathcal{I}_{\mathbf{A}^*} \subseteq \underline{\mathcal{I}}_{\mathcal{T}}(t_k) \backslash \tilde{\mathcal{I}}_{\mathcal{T}}$ be the set of indices like i such that $\mathrm{A}_i^{*\mathsf{T}} \mathbf{1}_m \geq 1$ where A_i^* is the i^{th} row of $\mathbf{A}^*$. One can see that $\mathbb{A}_{\mathcal{G}_{t_k}, \mathcal{I}_{\mathbf{A}^*}} \subseteq \tilde{\mathbb{A}}_{\mathcal{G}_{t_k}}$ and $\mathbf{A}^* \in \mathbb{A}_{\mathcal{G}_{t_k}, \mathcal{I}_{\mathbf{A}^*}}$. Hence

$$(\mathbf{x}^*, \mathbf{A}^*) \in \underset{(\mathbf{x},\mathbf{A}) \in \Omega_{t_k} \times \mathbb{A}_{\mathcal{G}_{t_k}, \mathcal{I}_{\mathbf{A}^*}}}{\operatorname{argmax}} \; \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}). \tag{4.60}$$

Define function $\bar{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}} : \mathcal{M}^{|\mathcal{I}_{\mathcal{V}}|} \times [0,1]^{|\underline{\mathcal{I}}_{\mathcal{T}}(t_k)| \times |\mathcal{I}_{\mathcal{V}}|} \to \mathbb{R}_{\geq 0}$ as

$$\bar{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}) = J_{\hat{\mathbf{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}) - \lambda \sum_{i \in \mathcal{I}_{\mathbf{A}^*} \backslash \tilde{\mathcal{I}}_{\mathcal{T}}} (\mathrm{A}_i^{\mathsf{T}} \mathbf{1}_m - 1), \tag{4.61}$$

where $\mathbf{A} = (\mathrm{A}_i)_{i \in \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}$. For any $(\mathbf{x}, \mathbf{A}) \in \Omega_{t_k} \times \mathbb{A}_{\mathcal{G}_{t_k}, \mathcal{I}_{\mathbf{A}^*}}$, one can simply verify that $\tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}) = \bar{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A})$. Therefore,

$$\underset{(\mathbf{x},\mathbf{A}) \in \Omega_{t_k} \times \mathbb{A}_{\mathcal{G}_{t_k}, \mathcal{I}_{\mathbf{A}^*}}}{\operatorname{argmax}} \; \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}) \;=\; \underset{(\mathbf{x},\mathbf{A}) \in \Omega_{t_k} \times \mathbb{A}_{\mathcal{G}_{t_k}, \mathcal{I}_{\mathbf{A}^*}}}{\operatorname{argmax}} \; \bar{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}). \tag{4.62}$$

Since $\bar{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}$ depends linearly on $\mathbf{A}$ and $\mathbb{A}_{\mathcal{G}_{t_k},\mathcal{I}_{\mathbf{A}^*}}$ is a polytope with etreme points belonging to $\mathbb{B}_{\mathcal{G}_{t_k},\mathcal{I}_{\mathbf{A}^*}}$, there exists $\mathbf{A}^{**} \in \mathbb{B}_{\mathcal{G}_{t_k},\mathcal{I}_{\mathbf{A}^*}}$ such that

$$\mathbf{A}^{**} \in \underset{\mathbf{A} \in \mathbb{A}_{\mathcal{G}_{t_k},\mathcal{I}_{\mathbf{A}^*}}}{\operatorname{argmax}} \ \bar{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^*, \mathbf{A}), \tag{4.63}$$

and therefore

$$\mathbf{A}^{**} \in \underset{\mathbf{A} \in \mathbb{A}_{\mathcal{G}_{t_k},\mathcal{I}_{\mathbf{A}^*}}}{\operatorname{argmax}} \ \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^*, \mathbf{A}). \tag{4.64}$$

Note that for any $\mathbf{x}^{**} \in \Omega_{t_k}$ such that

$$\mathbf{x}^{**} \in \underset{\mathbf{x} \in \Omega_{t_k}}{\operatorname{argmax}} \bar{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}^{**}), \tag{4.65}$$

one has

$$\mathbf{x}^{**} \in \underset{\mathbf{x} \in \Omega_{t_k}}{\operatorname{argmax}} \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}^{**}). \tag{4.66}$$

From (4.64) and (4.66), it yields that

$$\tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^*, \mathbf{A}^*) \leq \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^*, \mathbf{A}^{**}) \leq \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^{**}, \mathbf{A}^{**}). \tag{4.67}$$

Since $(\mathbf{x}^{**}, \mathbf{A}^{**}) \in \Omega_{t_k} \times \mathbb{A}_{\mathcal{G}_{t_k}, \mathcal{I}_{\mathbf{A}^*}}$, it concludes from (4.60) and (4.67) that $\tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^*, \mathbf{A}^*) = \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^{**}, \mathbf{A}^{**})$ and

$$(\mathbf{x}^{**}, \mathbf{A}^{**}) \in \underset{(\mathbf{x}, \mathbf{A}) \in \Omega_{t_k} \times \tilde{\mathbb{A}}_{\mathcal{G}_{t_k}}}{\operatorname{argmax}} \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}). \tag{4.68}$$

This shows that $(\mathbf{x}^{**}, \mathrm{A}^{**})$ is a solution of (4.38) with the entries of $\mathbf{A}^{**}$ being 0 or 1.

Let $\mathbf{A} = (a_{ij})$ be a $|\underline{\mathcal{I}}_{\mathcal{T}}(t_k)|$ by $|\mathcal{I}_{\mathcal{V}}|$ matrix with the property that a_{ij} is 1 only if i is the index according to $\emptyset_j$, for any $i \in \underline{\mathcal{I}}_{\mathcal{T}}(t_k)$ and any $j \in \mathcal{I}_{\mathcal{V}}$. Then, one can see that $(\mathbf{x}^k, \underline{\mathbf{A}}) \in \Omega_{t_k} \times \tilde{\mathbb{A}}_{\mathcal{G}_{t_k}}$ and $\tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^k, \underline{\mathbf{A}}) = 0$, and subsequently, $\tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^{**}, \mathbf{A}^{**}) \geq 0$. Set $\underline{\lambda} \in \mathbb{R}_{>0}$ as $1 + \sum_{j \in \mathcal{I}_{\mathcal{V}}} \sum_{i \in \mathcal{I}_{\mathcal{T},j}(t_k) \backslash \tilde{\mathcal{I}}_{\mathcal{T}}} d^k_{ij}$ and let $\lambda \geq \underline{\lambda}$. Then, one has $\mathrm{A}_i^{**\mathsf{T}} \mathbf{1}_m \leq 1$, where $\mathrm{A}_i^{**\mathsf{T}}$ denotes i^{th} row of $\mathbf{A}^{**}$, for any $i \in \mathcal{I}_{\mathcal{T}}(t_k)$. Since, if there exists $i \in \mathcal{I}_{\mathcal{T}}(t_k)$ such that $\mathrm{A}_i^{**\mathsf{T}} \mathbf{1}_m > 1$, as each entry of $\mathbf{A}^{**}$ belongs to the set $\{0, 1\}$, it follows that $\mathrm{A}_i^{**\mathsf{T}} \mathbf{1}_m \geq 2$, and subsequently, $p(\mathrm{A}_i^{**}) \geq 1$. From this, it yields that

$$\begin{aligned}
\tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^{**}, \mathbf{A}^{**}) &\leq \sum_{i \in \underline{\mathcal{I}}_{\mathcal{T}}(t_k) \backslash \tilde{\mathcal{I}}_{\mathcal{T}}} \sum_{j \in \mathcal{I}_{\mathcal{V},i}(t_k)} a_{ij} e^{-\bar{\gamma} \|\mathrm{x}_j - \hat{\mathrm{y}}_i^{k+1}\|} - \underline{\lambda} \\
&\leq \sum_{i \in \underline{\mathcal{I}}_{\mathcal{T}}(t_k) \backslash \tilde{\mathcal{I}}_{\mathcal{T}}} \sum_{j \in \mathcal{I}_{\mathcal{V},i}(t_k)} d^k_{ij} - \underline{\lambda} \\
&< 0,
\end{aligned}$$

which contradicts $\tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^{**}, \mathbf{A}^{**}) \geq 0$. Thus, for any $i \in \mathcal{I}_{\mathcal{T}}(t_k)$, one has $\mathrm{A}_i^{**\mathsf{T}} \mathbf{1}_m \leq 1$, i.e.

$$(\mathbf{x}^{**}, \mathrm{A}^{**}) \in \Omega_{t_k} \times \mathbb{A}_{\mathcal{G}_{t_k}}. \tag{4.69}$$

From (4.68), (4.69) and $\Omega_{t_k} \times \mathbb{A}_{\mathcal{G}_{t_k}} \subseteq \Omega_{t_k} \times \tilde{\mathbb{A}}_{\mathcal{G}_{t_k}}$, it yields that

$$(\mathbf{x}^{**}, \mathbf{A}^{**}) \in \operatorname*{argmax}_{(\mathbf{x},\mathbf{A}) \in \Omega_{t_k} \times \mathbb{A}_{\mathcal{G}_{t_k}}} \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \mathcal{I}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}). \tag{4.70}$$

This shows that $(\mathbf{x}^{**}, \mathrm{A}^{**})$ is a solution of (4.37).

With a similar discussion, one can show that if $\lambda \geq \underline{\lambda}$ then there exists $\bar{\mathrm{A}}^{**} \in \mathbb{B}_{\mathcal{G}_{t_k}}$ such that

$$\bar{\mathbf{A}}^{**} \in \operatorname*{argmax}_{\mathbf{A} \in \tilde{\mathbb{A}}_{\mathcal{G}_{t_k}}} P(\mathbf{A}). \tag{4.71}$$

It can be easily seen that for any $(\mathbf{x}, \mathbf{A}) \in \Omega_{t_k} \times \tilde{\mathbb{A}}_{\mathcal{G}_{t_k}}$, one has

$$J_{\hat{\mathbf{y}}^{k+1}_{\bullet \mathcal{I}_{\mathcal{T}}(t_k)}}(\mathbf{x}, \mathbf{A}) \leq P(\mathbf{A}), \tag{4.72}$$

and hence

$$J_{\hat{\mathbf{y}}^{k+1}_{\bullet \mathcal{I}_{\mathcal{T}}(t_k)}}(\mathbf{x}^{**}, \mathbf{A}^{**}) \leq P(\bar{\mathbf{A}}^{**}). \tag{4.73}$$

Now, consider the map $\mathbf{x} : \tilde{\mathbb{B}}_{\mathcal{G}_{t_k}} \to \Omega_{t_k}$ such that $\mathbf{x}(\mathbf{A}) = (\mathrm{x}_j(\mathbf{A}))_{j \in \mathcal{I}_{\mathcal{V}}}$ and for any $j \in \mathcal{I}_{\mathcal{V}}$,

$\mathrm{x}_j(\mathbf{A})$ is defined as

$$\mathrm{x}_j(\mathbf{A}) = \begin{cases} \mathrm{x}_j^k + \frac{\hat{\mathrm{y}}_{i_j}^{k+1} - \mathrm{x}_j^k}{\|\hat{\mathrm{y}}_{i_j}^{k+1} - \mathrm{x}_j^k\|} u_{\max} H_k, & \text{if } i_j \notin \tilde{\mathcal{I}} \\ \mathrm{x}_j^k, & \text{if } i_j \in \tilde{\mathcal{I}} \end{cases} \quad (4.74)$$

where $i_j \in \underline{\mathcal{I}}_{\mathcal{T}}(t_k)$ is the index of the row that j^{th} column of $\mathbf{A}$ is 1 in that row. Hence, for any $\mathbf{A} \in \tilde{\mathbb{B}}_{\mathcal{G}_{t_k}}$, it yields that $\tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}(\mathbf{A}), \mathbf{A}) = P(\mathbf{A})$. Thus, according to the definition of $(\mathbf{x}^{**}, \mathbf{A}^{**})$, it can be noticed that

$$\tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^{**}, \mathbf{A}^{**}) = \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}(\mathbf{A}^{**}), \mathbf{A}^{**}) = P(\mathbf{A}^{**}), \quad (4.75)$$

and subsequently,

$$P(\bar{\mathbf{A}}^{**}) = \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}(\bar{\mathbf{A}}^{**}), \bar{\mathbf{A}}^{**}) \leq \tilde{J}_{\hat{\mathrm{y}}^{k+1}_{\bullet \underline{\mathcal{I}}_{\mathcal{T}}(t_k)}}(\mathbf{x}^{**}, \mathbf{A}^{**}) = P(\mathbf{A}^{**}). \quad (4.76)$$

Considering equations (4.73), (4.75) and (4.76), it follows that $P(\bar{\mathbf{A}}^{**}) = P(\mathbf{A}^{**})$, i.e.

$$\mathbf{A}^{**} \in \underset{\mathbf{A} \in \tilde{\mathbb{A}}_{\mathcal{G}_{t_k}}}{\operatorname{argmax}} P(\mathbf{A}). \quad (4.77)$$

This shows that $\mathbf{A}^{**}$ is a maximizer for the potential function (4.41) and hence a pure Nash equilibrium for $\mathfrak{G}_k$.

Chapter 5

Cooperative Receding Horizon Control of Double Integrator Vehicles for Multi-Target Interception

In this chapter, the cooperative multi-target interception problem in an uncertain environment with double-integrators vehicles is investigated. A time-discounting reward function is defined for each target which can be collected only if it is visited by a vehicle. This function is used to formulate the problem as an optimization problem which aims to maximize the expected reward collectible from the set of available targets in the mission space. A cooperative receding horizon controller is designed to solve the problem based on an estimate of the future position of every targets with the available information. It is shown that a solution for this optimization problem exists, and that the vehicles visit the targets in finite time. The effectiveness of the proposed algorithm is demonstrated by simulation.

5.1 Notations

Throughout the paper, the set of real numbers and the set of non-negative real numbers are denoted by $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$, respectively. Also, let $\mathbb{N}$ and $\mathbb{N}_n$ denote respectively the set of natural numbers and the set of natural numbers less than or equal to n. For a given set A and a subset of it B, the indicator function of B, denoted by $\mathbf{1}_B$, is a function from A to $\{0, 1\}$, which is non-zero only when its argument belongs to the set B. For any index set $\mathcal{I}$, the notation $A^{\mathcal{I}}$ represents the set of points like $(a_i)_{i \in \mathcal{I}}$ whose entries belong to A. In the case when I is the set $\mathbb{N}_n$, the set $A^{\mathcal{I}}$ is simply denoted by A^n. Let $\mathcal{J}$ be a non-empty subset of $\mathcal{I}$. For any point $\mathrm{a} \in A^{\mathcal{I}}$, the term $\mathrm{a}_{\bullet \mathcal{J}}$ represents a point in $A^{\mathcal{J}}$, obtained by eliminating the entries with indices not listed in $\mathcal{J}$. The d-dimensional Euclidean space is denoted by $\mathbb{R}^d$.

Let n be a natural number. Then, $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. Also, $\mathbf{0}_n$ and $\mathbf{1}_n$ represent all-zero and all-one vectors in $\mathbb{R}^n$, respectively. For any vector $\mathrm{a}, \mathrm{b} \in \mathbb{R}^d$, it is said that $\mathrm{a} \geq \mathrm{b}$ only when $\mathrm{a} - \mathrm{b} \in \mathbb{R}^n_{\geq 0}$, i.e., all entries of $\mathrm{a} - \mathrm{b}$ are non-negative. Let x be a point in $\mathbb{R}^d$ and r be a scalar in $\mathbb{R}_{\geq 0}$. Then, $\mathscr{B}(\mathrm{x}, r)$ denotes the closed ball in $\mathbb{R}^d$ with radius r centered at x, i.e.

$$\mathscr{B}(\mathrm{x}, r) = \{\mathrm{y} \in \mathbb{R}^d \mid \|\mathrm{x} - \mathrm{y}\| \leq r\}. \tag{5.1}$$

Let $\bar{q}$ and $\bar{u}$ be scalars in $\mathbb{R}_{\geq 0}$, and I be an interval in $\mathbb{R}$. Denote by $C_I^{\wp}(\mathbb{R}^d)$ the set of piecewise continuous vector-valued functions defined over I with values in $\mathbb{R}^d$. Accordingly, define $\mathscr{U}_I(\bar{u})$ as a set of functions in $C_I^{\wp}(\mathbb{R}^d)$ like u such that $\sup_{t \in I} \|\mathrm{u}(t)\| \leq \bar{u}$. In the case where $\bar{u}$ is known from the context, the arguments will be omitted for brevity.

5.2 Problem Formulation

Let the mission space, denoted by $\mathcal{M}$, be a closed convex subset of d-dimensional Euclidean space, and $\mathcal{I}_{\mathcal{V}} = \mathbb{N}_{|\mathcal{I}_{\mathcal{V}}|}$ be the set of indices for a finite number of vehicles inside $\mathcal{M}$. For any $j \in \mathcal{I}_{\mathcal{V}}$, let $\mathrm{p}_j(t) \in \mathbb{R}^d$ and $\mathrm{q}_j(t) \in \mathbb{R}^d$ represent the position vector and velocity vector of vehicle j, at time $t \in \mathbb{R}_{\geq 0}$, whose dynamics is described by

$$\begin{cases} \dot{\mathrm{p}}_j(t) = \mathrm{q}_j(t), \\ \dot{\mathrm{q}}_j(t) = \mathrm{u}_j(t), \end{cases} \tag{5.2}$$

where u_j belongs to the set of *admissible controls*, denoted by $\mathscr{U}$ and defined here as $\mathscr{U}_I(u_{\max})$ where $u_{\max}$ is the bounds of acceleration for the vehicles. For any $j \in \mathcal{I}_\mathcal{V}$, denote by x_j the state vector of vehicle j, which is defined as $[\mathrm{p}_j^\intercal, \mathrm{q}_j^\intercal]^\intercal$ and belongs to the set $\mathcal{M} \times \mathbb{R}^d$. Accordingly, $\mathbf{x} = (\mathrm{x})_{j\in\mathcal{I}_\mathcal{V}}$ is the state vector of the entire system. The set $\mathcal{M} \times \mathbb{R}^d$ is the state space for each of the vehicles and $\mathcal{X} = \times_{j\in\mathcal{I}_\mathcal{V}}(\mathcal{M} \times \mathbb{R}^d)$ is the state space for the entire system.

Remark 12. *It is to be noted that* $\mathrm{u}_j(t) = u_j(t)\mathrm{d}_j(t)$ *for some piecewise continuous functions* u_j *and* d_j*, where* $\mathrm{d}_j(t) \in \mathbb{S}^{d-1} = \{\mathrm{d} \in \mathbb{R}^d; \|\mathrm{d}\| = 1\}$ *is the control input for the direction of the acceleration vector and* $u_j(t) \in [0, u_{\max}]$ *is the control input for its magnitude, for any* $j \in \mathcal{I}_\mathcal{V}$ *and any* $t \in \mathbb{R}_{\geq 0}$.

Let $\mathcal{I}_\mathcal{T} = \mathbb{N}_{|\mathcal{I}_\mathcal{T}|}$ be a finite set of natural numbers representing indices of a non-zero finite number of targets sequentially arriving in the mission space. Assume that the mission starts at $t = 0$, and let $n_0 \in \{0\} \cup \mathbb{N}_{|\mathcal{I}_\mathcal{T}|}$ be the number of targets in the mission space initially. Let also T_1 be the arrival time of the first target, and $\{T_i\}_{i=2}^{|\mathcal{I}_\mathcal{T}|}$ be a finite sequence of non-negative real scalars representing targets inter-arrival times, i.e., the time between consecutive targets' arrival. Note that if $n_0 > 0$, then for any $1 \leq i \leq n_0$, one has $T_i = 0$. For any $i \in \mathbb{N}_{|\mathcal{I}_\mathcal{T}|}$, one can define the arrival time of the i^{th} target as $\check{\tau}_i = \sum_{j=1}^{i} T_j$, and also the set of indices of targets arrived up to time instant t, denoted by $\check{\mathcal{I}}_\mathcal{T}(t)$, as

$$\check{\mathcal{I}}_\mathcal{T}(t) := \{i \in \mathcal{I}_\mathcal{T}\ ;\ \check{\tau}_i \leq t\}. \tag{5.3}$$

It is worth noting that $\{\check{\tau}_i\}_{i\in\mathcal{I}_\mathcal{T}}$ is an increasing finite sequence. Besides the arrival time of targets, one can define a sequence of vectors $\{\mathrm{r}_i\}_{i\in\mathcal{I}_\mathcal{T}}$, belonging to $\mathcal{M}$, as the initial positions of targets in the mission space as they arrive. The arrival times and initial positions of targets are not known *a priori*. More precisely, at any time $t < \check{\tau}_i$, none of the vehicles has the information of $\check{\tau}_i$ and r_i. In other words, for any $i \in \mathcal{I}_\mathcal{T}$, the i^{th} target arrives in the mission space at an *a priori* unknown time $\check{\tau}_i$ and in an *a priori* unknown point r_i. In addition, the vehicle moves on an *a priori* unknown trajectory, denoted by $\mathrm{r}_i(t)$.

Definition 12. *For any $i \in \mathcal{I}_\mathcal{T}$ and $j \in \mathcal{I}_\mathcal{V}$. and a prescribed positive scalar d_{ij}, it is said that the j^{th} vehicle visits the i^{th} target at time t, if $\|\mathrm{p}_j(t) - \mathrm{r}_i(t)\| \leq d_{ij}$.*

Along with Definition 12 and for any $i \in \mathcal{I}_\mathcal{T}$, one can define $\hat{\tau}_i \in \bar{\mathbb{R}}_{\geq 0} = [0, \infty]$ as the first time that target i is visited by one of the vehicles, i.e.

$$\hat{\tau}_i = \inf\{t \in \mathbb{R}_{\geq 0} \; ; \; \min_{j\in\mathcal{I}_\mathcal{V}} \left(\|\mathrm{p}_j(t) - \mathrm{r}_i(t)\| - d_{ij}\right) \leq 0\}. \tag{5.4}$$

Note that $\hat{\tau}_i = \infty$ if and only if none of the vehicles visits target i. The set of indices of targets visited up to time t, denoted by $\hat{\mathcal{I}}_\mathcal{T}(t)$, is defined as

$$\hat{\mathcal{I}}_\mathcal{T}(t) = \{i \in \mathcal{I}_\mathcal{T} \mid \hat{\tau}_i \leq t\}. \tag{5.5}$$

One can also define the set of indices of targets arrived in the mission space but not visited up to time t, as

$$\mathcal{I}_{\mathcal{T}}(t) = \check{\mathcal{I}}_{\mathcal{T}}(t) \backslash \hat{\mathcal{I}}_{\mathcal{T}}(t) = \{ i \in \mathcal{I}_{\mathcal{T}} \ ; \ \check{\tau}_i \leq t < \hat{\tau}_i \}. \tag{5.6}$$

For any $i \in \mathcal{I}_{\mathcal{T}}$, the trajectory of target i is a C^2 curve in the mission space $\mathcal{M}$, defined by $\mathrm{r}_i : [\check{\tau}_i, \hat{\tau}_i] \rightarrow \mathcal{M}$, satisfying the following two geometric conditions where one describes the global behavior of trajectories of targets and the other one describes the local behavior of trajectories of targets.

Assumption 8. *(Global Geometric Condition) For any $i \in \mathcal{I}_{\mathcal{T}}$ and any $\tau \in [\check{\tau}_i, \hat{\tau}_i]$, $\mathrm{r}_i(\tau) \in \mathcal{M}$.*

Global geometric condition guarantees that once a target arrives, it will remain inside the mission space until the end of the mission. Note that the property stated in Assumption 8 depends both on trajectories of targets and also on the geometry of the mission space. For example, in the case where $\mathcal{M}$ is the whole d-dimensional Euclidean space, one can verify that the global geometric condition is satisfied automatically.

Assumption 9. *(Local Geometric Condition) For any $i \in \mathcal{I}_{\mathcal{T}}$, $\mathrm{r}_i : [\check{\tau}_i, \hat{\tau}_i] \rightarrow \mathcal{M}$ is a C^2 function, i.e. r_i is two times continuously differentiable. Also, there exist non-negative scalars $\tilde{a}_i, \tilde{\jmath}_i, \tilde{c}_i$ such that for any $i \in \mathcal{I}_{\mathcal{T}}$, $\tau \in [\check{\tau}_i, \hat{\tau}_i]$ and $t \in (\tau, \hat{\tau}_i]$, one has*

$$\| \frac{\mathrm{d}^2}{\mathrm{d}t^2} \mathrm{r}_i(\tau) \| \leq \tilde{a}_i, \qquad \sup_{\tilde{\tau} \in (\tau, \hat{\tau}_i]} \| \mathrm{j}_i(\tilde{\tau}, \tau) \| \leq \tilde{\jmath}_i, \tag{5.7}$$

and

$$\|\mathrm{j}_i(t,\tau)\| \le \tilde{c}_i|t-\tau|, \tag{5.8}$$

where $\mathrm{j}_i(t,\tau)$ *is a* C^2 *function satisfying the following equality*

$$\mathrm{r}_i(t) = \mathrm{r}_i(\tau) + (t-\tau)\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{r}_i(\tau) + \frac{1}{2!}(t-\tau)^2\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{r}_i(\tau) + \frac{1}{3!}(t-\tau)^3\mathrm{j}_i(t,\tau). \tag{5.9}$$

Assumption 10. *The position, velocity and acceleration vectors of any current target (targets that have arrived but not visited yet) are available at the beginning of each time horizon (i.e., at time instant* τ *in (5.9)).*

For any $\tau \in \mathbb{R}_{\ge 0}$ and any $i \in \mathcal{I}_{\mathcal{T}}(\tau)$, define $\mathrm{y}_i(\tau)$ as the vector of available information of target i at time instant τ, i.e.

$$\mathrm{y}_i(\tau) = \big(\mathrm{r}_i(\tau), \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{r}_i(\tau), \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{r}_i(\tau)\big). \tag{5.10}$$

This information vector belongs to the information space of target i, which is defined as $\mathcal{Y}_i = \mathcal{M} \times \mathbb{R}^d \times \mathscr{B}(\mathbf{0}_d, \tilde{a}_i)$. Accordingly, one can define the *information vector of targets* at time τ as

$$\mathbf{y}(\tau) = \big(\mathrm{r}_i(\tau), \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{r}_i(\tau), \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{r}_i(\tau)\big)_{i \in \mathcal{I}_{\mathcal{T}}(\tau)}, \tag{5.11}$$

and also, the *information space of targets* as

$$\mathcal{Y}_\tau = \bigtimes_{i \in \mathcal{I}_\mathcal{T}(\tau)} \big(\mathcal{M} \times \mathbb{R}^d \times \mathscr{B}(\mathbf{0}_d, \tilde{a}_i)\big). \tag{5.12}$$

Considering Assumption 10, for any $\tau \in \mathbb{R}_{\geq 0}$ and any $i \in \mathcal{I}_\mathcal{T}(\tau)$, the position of target i can be estimated at any future instant within the time horizon of its presence in the mission space using the information available at time τ. Denote this estimate by $\hat{\mathrm{r}}_i(\cdot)$, and describe it by

$$\hat{\mathrm{r}}_i(t) = \mathrm{r}_i(\tau) + (t-\tau)\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{r}_i(\tau) + \frac{1}{2!}(t-\tau)^2\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{r}_i(\tau), \tag{5.13}$$

where $t \in [\tau, \hat{\tau}_i]$.

With respect to each target, a *task* is defined which is completed if the target is visited by one of the vehicles. By slight abuse of notation, denote by $\mathcal{I}_\mathcal{T}$, $\check{\mathcal{I}}_\mathcal{T}(t)$, $\hat{\mathcal{I}}_\mathcal{T}(t)$ and $\mathcal{I}_\mathcal{T}(t)$, the total set of tasks, the set of tasks started by time t, the set of tasks accomplished by time t, and the set of tasks in progress at time t, respectively. Subsequently, the *mission* is to accomplish all of the tasks in finite time. Here, it is desired to obtain a near-optimal cooperative algorithm to accomplish the mission in the presence of uncertainties and limited information.

5.3 Cooperative Receding Horizon Scheme

As an incentive for the vehicles to accomplish the tasks (i.e. visit the targets), let a time-decreasing reward be assigned to each target, which can be collected only if the vehicle completes the corresponding task and visits the target. Vehicles intend to maximize the total collected reward, which entails cooperation to minimize the visit time. Toward this goal, each vehicle should decide upon its next immediate target in a cooperative manner, during the decision-making process, and subsequently plan its own path. Due to uncertainties in the environment and changes in the required information, the cooperative decision-making and path planning process should be performed iteratively. At the beginning of each iteration, the vehicles calculate their control inputs based on the tasks and their corresponding rewards, such that their estimation of the total collected reward is maximized.

5.3.1 Structure of Reward Functions

For any $i \in \mathcal{I}_{\mathcal{T}}$, let $\mathscr{R}_i$ be the *initial reward* considered for the task corresponding to the i^{th} target at its arrival moment. In order to take into account the reward loss over time, define a continuous decreasing function $\rho_i : [\check{\tau}_i, \hat{\tau}_i] \to [0, 1]$, called *discount function*, and form the *reward function* as $\mathscr{R}_i \rho_i$. Assuming that $\rho_i(\check{\tau}_i) = 1$, the reward function satisfies the desired properties discussed earlier. By properly selecting the initial rewards and discount functions amongst their possible candidates, one can model aspects such as scheduling, time priorities and deadlines. For example, in the case that there is no final deadline for visiting the target i, one can consider the discount function as follows

$$\rho_i(t) = e^{-\gamma_i(t-\check{\tau}_i)}, \quad \forall\, i \in \mathcal{I}_{\mathcal{T}}, \tag{5.14}$$

where $\gamma_i \in \mathbb{R}_{>0}$ is the *reward discount rate parameter* for the target i. Also, in the case that there is a final hard deadline for task i, denoted by $t^i_f \in \mathbb{R}_{\geq 0}$, one may consider the following discount function

$$\rho_i(t) = \max\{1 - \frac{t - \check{\tau}_i}{t^i_f - \check{\tau}_i}, 0\}, \quad i \in \mathcal{I}_{\mathcal{T}}. \tag{5.15}$$

Moreover, one may consider finite number of soft deadlines after which the corresponding target is not as interesting as it was before the deadline. For this case, one may choose the discount functions as continuous piecewise-defined functions formed by some other discount functions as its sub-functions, i.e.

$$\rho_i(t) = \begin{cases} \rho_{i,0}(t), & \text{if } t \in [\check{\tau}_i, \mathcal{D}_{i1}), \\ \rho_{i,1}(t), & \text{if } t \in [\mathcal{D}_{i1}, \mathcal{D}_{i2}), \\ \vdots & \vdots \\ \rho_{i,d^i}(t), & \text{if } t \in [\mathcal{D}_{id^i}, \hat{\tau}_i], \end{cases} \tag{5.16}$$

where $\{\mathcal{D}_{i\mathscr{d}}\}_{\mathscr{d}=1}^{\mathscr{d}^i}$ are the soft deadlines and $\{\rho_{i,\mathscr{d}}(\cdot)\}_{\mathscr{d}=1}^{\mathscr{d}^i}$ are the discount sub-functions. Since $\rho_i(t)$ is a continuous function, it is required that for any $d \in \mathbb{N}_{\mathscr{d}^i}$, one has $\lim_{t\to\mathcal{D}_{i\mathscr{d}}} \rho_{i,\mathscr{d}-1}(\mathcal{D}_{i\mathscr{d}}) = \rho_{i,\mathscr{d}}(\mathcal{D}_{i\mathscr{d}})$. Note that some of these soft deadlines may be considered based on the unpredicted events occurring in the mission and thus, they are *a priori* unknown.

5.3.2 The Minimum Reaching Time and The Maximum Reward Estimation

Let $\{t_k\}_{k=1}^{k_{\max}} \in [0, \hat{\tau}]$ denote the time instants when the iterative decision-making procedure is supposed to be performed where $k_{\max} \in \mathbb{N} \cup \{\infty\}$ represents the number of iterations. Note that it is implicitly assumed that $t_1 = 0$ and the sequence $(t_k)_{k=1}^{k_{\max}}$ is a strictly increasing sequence. Accordingly, for any $k \in \mathbb{N}$ such that $k < k_{\max}$, one can define the k^{th} time-interval of procedure as $I_k = [t_k, t_{k+1})$.

For any $j \in \mathcal{I}_{\mathcal{V}}$, from (5.2), one can represent the dynamic of vehicle j in matrix form as follows

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{x}_j = \mathbf{A}_d \mathrm{x}_j + \mathbf{B}_d \mathrm{u}_j \tag{5.17}$$

where $\mathbf{A}_d$ and $\mathbf{B}_d$ are $2d$ by $2d$ matrices defined as

$$\mathbf{A}_d = \begin{pmatrix} \mathbf{0}_d & \mathbf{I}_d \\ \mathbf{0}_d & \mathbf{0}_d \end{pmatrix}, \mathbf{B}_d = \begin{pmatrix} \mathbf{0}_d \\ \mathbf{I}_d \end{pmatrix}. \tag{5.18}$$

Similarly, the dynamics of the all system is derived in matrix form as following

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = (\mathbf{I}_m \otimes \mathbf{A}_d)\mathbf{x} + (\mathbf{I}_m \otimes \mathbf{B}_d)\mathbf{u}. \tag{5.19}$$

For any $j \in \mathcal{I}_{\mathcal{V}}$, let $\mathrm{u}_j^k(\cdot)$ be a function in $\mathscr{U}_{I_k}$, the set of admissible controls defined over I_k the control input applied by the vehicle j for time interval $[t_k, t_{k+1})$. Subsequently, let $\mathbf{u}^k$ denote the vector of all control inputs $(\mathrm{u}_j^k)_{j\in\mathcal{I}_{\mathcal{V}}}$. Under these control inputs, the state of vehicle j at time instant t_{k+1} is obtained as

$$\mathrm{x}_j(t_{k+1}) = e^{\mathbf{A}_d(t_{k+1}-t_k)}\mathrm{x}_j(t_k) + \int_{t_k}^{t_{k+1}} e^{\mathbf{A}_d(t_{k+1}-s)}\mathbf{B}_d\mathrm{u}_j^k(s)\mathrm{d}s. \tag{5.20}$$

Now, let $i \in \mathcal{I}_{\mathcal{T}}(t_k)$ be the index of an arbitrary existing target. Let τ_{ij}^k be the time estimate when the vehicle j can reach target i, based on the information given at time instant t_k and the control input $\mathrm{u}_j^k(\cdot)$ applied by the vehicle j for time interval $[t_k, t_{k+1})$. Denote r_i^k, v_i^k and a_i^k as $\mathrm{r}_i(t_k)$, $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{r}_i(t_k)$ and $\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{r}_i(t_k)$, respectively. From these, one can model the trajectory of target i, for $t \geq t_k$ as

$$\hat{\mathrm{r}}_i(t) = \mathrm{r}_i^k + (t - t_k)\mathrm{v}_i^k + \frac{1}{2!}(t - t_k)^2\mathrm{a}_i^k. \tag{5.21}$$

Consequently, one can obtain

$$\begin{cases} \hat{\mathrm{r}}_i^{k+1} &= \mathrm{r}_i^k + (t_{k+1} - t_k)\mathrm{v}_i^k + \frac{1}{2!}(t_{k+1} - t_k)^2 \mathrm{a}_i^k, \\ \hat{\mathrm{v}}_i^{k+1} &= \mathrm{v}_i^k + (t_{k+1} - t_k)\mathrm{a}_i^k, \\ \hat{\mathrm{a}}_i^{k+1} &= \mathrm{a}_i^k, \end{cases} \tag{5.22}$$

where $\hat{\mathrm{r}}_i^{k+1}$, $\hat{\mathrm{v}}_i^{k+1}$ and $\hat{\mathrm{a}}_i^{k+1}$ are the prediction of the position, velocity and acceleration of target i at t_{k+1}, respectively, based on the information given at time instant t_k. Having these predictions, one can obtain the following theorem based on minimum time optimal control theory.

Theorem 10. *Consider the vehicle $j \in \mathcal{I}_\mathcal{V}$, the time instant t_k, the target $i \in \mathcal{I}_\mathcal{T}(t_k)$ and the trajectory model given in* (5.21) *for the target i. Let the control input $\mathrm{u}_j^k(\cdot)$ be applied for time interval I_k. Also, consider the following equation*

$$\begin{cases} \frac{1}{2}\mathrm{u}_{ij}\bar{\tau}_{ij}^2 &= \frac{1}{2}\hat{\mathrm{a}}_i^{k+1}\bar{\tau}_{ij}^2 + (\hat{\mathrm{v}}_i^{k+1} - \mathrm{q}_j(t_{k+1}))\bar{\tau}_{ij} + (\hat{\mathrm{r}}_i^{k+1} - \mathrm{p}_j(t_{k+1})) \\ \|\mathrm{u}_{ij}\| &= u_{\max}, \end{cases} \tag{5.23}$$

where $\hat{\mathrm{r}}_i^{k+1}$, $\hat{\mathrm{v}}_i^{k+1}$ and $\hat{\mathrm{a}}_i^{k+1}$ are the predictions given in (5.22) *and also, $\mathrm{p}_j(t_{k+1})$ and $\mathrm{q}_j(t_{k+1})$ are the position and velocity of vehicle j, respectively, given the control input $\mathrm{u}_j^k(\cdot)$ is applied for time interval I_k. Then, equation* (5.23) *has a solution for $\bar{\tau}_{ij}$ in $\mathbb{R}_{\geq 0}$, and subsequently, a corresponding solution for u_{ij}. Also, if one has that $t_{k+1} - t_k \leq z$*

where z is smallest positive solution of following equations

$$\frac{1}{2}(u_{\max} + \tilde{a})z^2 = \|(\mathrm{q}_j(t_k) - \mathrm{v}_i^k)z + \mathrm{p}_j(t_k) - \mathrm{r}_i^k\|, \tag{5.24}$$

then $\tau_{ij}^k = t_{k+1} + \bar{\tau}_{ij}$, where $\bar{\tau}_{ij}$ is the smallest non-negative solution of equation (5.23).

Corollary 3. *Let the conditions in Theorem 10 hold. Then the maximum reward which vehicle j can collect from target i, assuming that the control input $\mathrm{u}_j^k(\cdot)$ is applied for the time interval I_k, can be estimated as $\mathcal{R}_i\rho_i(\tau_{ij}^k)$ where τ_{ij}^k is the estimation of the reaching time introduced in Theorem 10.*

The Theorem 10 and Corollary 3 say that based on the given information, the vehicles can estimate the minimum reaching times and subsequently, the final maximum rewards where each of them can extract from each of the targets. However, in order to maximize the total collected reward, the vehicles are required to cooperate in an appropriate manner. The structure of this cooperation is discussed in the sequel.

5.3.3 Structure of Cooperation Strategy

In each step of decision-making, once the vehicles estimate the final maximum reward of each target using the given information of the targets and vehicles, each of them is required to decide upon its next immediate target. Based on the possible differences in the value of estimated final maximum rewards of different targets, vehicles may have different levels of interest in the tasks. However, they should cooperate in order to maximize the final total rewards collected from all the targets. In this regard, a cooperation strategy is required which is discussed here. Consider the step of decision making corresponding to the time instant t_k. For any $i \in \mathcal{I}_{\mathcal{T}}(t_k)$ and $j \in \mathcal{I}_{\mathcal{V}}$, an *assignment* of task i to vehicle j is characterized as a real scalar in $[0, 1]$, denoted by π_{ij}^k, which reflects the amount of interest of vehicle j in being target i assigned to it during the time interval I_k. Also, denote by $\mathbf{\Pi}^k$ the *assignments matrix* which is defined as $(\pi_{ij}^k)_{i \in \mathcal{I}_{\mathcal{T}}(t_k), j \in \mathcal{I}_{\mathcal{V}}}$. It is expected that the value of assignment π_{ij}^k depends implicitly on the information given at time instant t_k via the estimation of final maximum rewards and also the cooperation policy constraints. More precisely, for any $i \in \mathcal{I}_{\mathcal{T}}(t_k)$ and $j \in \mathcal{I}_{\mathcal{V}}$, the assignment π_{ij}^k is a function of the form $\pi_{ij}^k : \mathcal{X} \times \mathcal{Y}_{t_k} \to [0, 1]$ where $\mathcal{X} \times \mathcal{Y}_{t_k}$ is the information space at time instant t_k. The proper assignments are required to have some desired structures reflecting cooperation policy constraints which are discussed in the sequel.

If t_k is a time instant such that $\mathcal{I}_{\mathcal{T}}(t_k) = \emptyset$, there is no target in the mission space and no issue for cooperation and assignment. Therefore, let t_k be a time instant at which $\mathcal{I}_{\mathcal{T}}(t_k) \neq \emptyset$. Since the vehicles are required to consider all the current tasks, it is expected that for each vehicle the sum of its assignments be equal to one, i.e.

$$\sum_{i \in \mathcal{I}_{\mathcal{T}}(t_k)} \pi_{ij}^k(\mathbf{x}^k, \mathbf{y}^k) = 1, \quad \forall j \in \mathcal{I}_{\mathcal{V}}, \tag{5.25}$$

where $\mathbf{x}^k$ and $\mathbf{y}^k$ are vectors for states of vehicles and the available information of targets, respectively, at time instant t_k. Also, in the case that the number of current tasks is at least equal to the number of vehicles, it is reasonable to manage the resources efficiently to accomplish as many tasks as possible by acting cautiously. Hence, it is required to under-assign the targets to the vehicles, i.e.

$$\sum_{j \in \mathcal{I}_\mathcal{V}} \pi_{ij}^k(\mathbf{x}^k, \mathbf{y}^k) \leq 1, \quad \forall i \in \mathcal{I}_\mathcal{T}(t_k). \tag{5.26}$$

Similarly, in the case that the number of vehicles is at least equal to the number of current tasks, according to the possible uncertainties in the environment, it is expected to increase the chance of collecting more amounts of reward by acting generously. Therefore, it is required to over-assign the targets to the vehicles, i.e.

$$\sum_{j \in \mathcal{I}_\mathcal{V}} \pi_{ij}^k(\mathbf{x}^k, \mathbf{y}^k) \geq 1, \quad \forall i \in \mathcal{I}_\mathcal{T}(t_k). \tag{5.27}$$

Remark 13. *It can be shown that the inequalities in equations* (5.26) *and* (5.27) *turn to equalities when* $|\mathcal{I}_\mathcal{V}| = |\mathcal{I}_\mathcal{T}(t_k)|$.

Equations (5.25), (5.26) and (5.27) introduce a set of constraints that should be satisfied by any desired assignment. More precisely, if one defines the set $\mathbb{P}^{n\times m}$ as

$$\mathbb{P}^{n\times m} = \{\mathbf{\Pi} \in [0,1]^{n\times m}; \mathbf{\Pi}^\mathsf{T}\mathbf{1}_n = \mathbf{1}_m, m \geq n \Rightarrow \mathbf{\Pi}\mathbf{1}_m \geq \mathbf{1}_n, m \leq n \Rightarrow \mathbf{\Pi}\mathbf{1}_m \leq \mathbf{1}_n\}, \tag{5.28}$$

for any $n, m \in \mathbb{N}$, then the assignment matrix $\mathbf{\Pi}^k(\mathbf{x}^k, \mathbf{y}^k)$ is required to belong to the set $\mathscr{P}_{\mathcal{I}_\mathcal{T}(t_k),\mathcal{I}_\mathcal{V}}$ which is defined as $\mathscr{P}_{\mathcal{I}_\mathcal{T}(t_k),\mathcal{I}_\mathcal{V}} = \{\mathbf{\Pi} : \mathcal{X} \times \mathcal{Y}_{t_k} \to \mathbb{P}^{|\mathcal{I}_\mathcal{T}(t_k)| \times |\mathcal{I}_\mathcal{V}|}\}$.

5.3.4 Cooperative Receding Horizon Controller

The cooperative receding horizon (CRH) controller performs the iterative procedure of cooperative decision-making and path planning. The controller generates the control inputs for each vehicle as well as the matrix of optimal assignments such that the vehicles collect maximum possible rewards. Toward this goal, an estimation of the remaining collectible rewards is given as a payoff function in an optimization problem, at any time instant t_k, and the solution of the problem is obtained. The payoff function depends on the control inputs and assignments for the current time step. The constraints in the optimization problem and also the payoff function are mainly based on the information given at time instant t_k. The solution of the problem provides the optimal control input $\mathbf{u}^k$.

Let t_k be a time instant such that $\mathcal{I}_\mathcal{T}(t_k) \neq \emptyset$, and $(\mathrm{u}_j^k(\cdot))_{j \in \mathcal{I}_\mathcal{V}}$ be the control inputs applied to the vehicles for time period $I_k = [t_k, t_{k+1})$. Therefore, the states of the vehicles at time instant t_{k+1}, the vector $\mathbf{x}^{k+1}$, is derived as in (5.20). Moreover, equation (5.22) provides the vector of predictions of position, velocity and acceleration of targets at t_{k+1}, the vector $\hat{\mathbf{y}}^{k+1}$. Consider the estimation of the final maximum rewards introduced in Corollary 3, i.e. $\mathcal{R}_i \rho_i(\tau_{ij}^k(\mathbf{u}^k, t_k))$ which is the estimation of the maximum value of reward that vehicle j expects at time instant t_{k+1} to collect from target i given that the control input $\mathrm{u}_j^k(\cdot)$ is applied for time interval $I_k = [t_k, t_{k+1})$, for any $i \in \mathcal{I}_{\mathcal{T}}(t_k)$ and $j \in \mathcal{I}_{\mathcal{V}}$. Also, consider the *expected optimal assignment matrix* for time instant t_{k+1}, denoted by $\tilde{\mathbf{\Pi}}^{k+1}$, defined as the optimal assignment matrix. This matrix is determined based on the state vector $\mathbf{x}^{k+1}$ as well as the prediction vector $\hat{\mathbf{y}}^{k+1}$ which itself depends on the information provided at t_k. Accordingly, one can say that the expected optimal assignment matrix is a function of control input $\mathbf{u}^k$ and time instant t_k, i.e. $\tilde{\mathbf{\Pi}}^{k+1} = \tilde{\mathbf{\Pi}}^{k+1}(\mathbf{u}^k, t_k)$. Given the estimation of rewards, the expected optimal assignment matrix, the states of vehicles and the vector of predictions regarding the targets, all at time instant t_{k+1}, one can estimate at t_{k+1}, the maximum reward the team expects to collect until the end of mission. This total expected reward is denoted by $\mathfrak{R}^{k+1}(\mathbf{u}^k, t_k)$ and formulated as following

$$\mathfrak{R}^{k+1}(\mathbf{u}^k, t_k) = \sum_{i \in \mathcal{I}_{\mathcal{T}}(t_k)} \sum_{j \in \mathcal{I}_{\mathcal{V}}} \mathcal{R}_i \rho_i(\tau_{ij}^k(\mathbf{u}^k, t_k)) \tilde{\pi}_{ij}^k(\mathbf{u}^k, t_k), \tag{5.29}$$

where $\tilde{\pi}_{ij}^k(\mathbf{u}^k, t_k)$ is the entry of matrix $\tilde{\mathbf{\Pi}}^{k+1}(\mathbf{u}^k, t_k)$ in row i and column j, for any $i \in \mathcal{I}_{\mathcal{T}}(t_k)$ and $j \in \mathcal{I}_{\mathcal{V}}$.

Now, let P^k be the optimization problem for CRH controller corresponding to k^{th} step. According to the discussion above, one has

$$\mathrm{P}^k : \begin{cases} \max & \mathfrak{R}^{k+1}(\mathbf{u}^k, t_k) \\ \text{s.t.} & \tilde{\mathbf{\Pi}}(\mathbf{u}^k, t_k) \in \mathscr{P}^k, \mathbf{u}^k \in \mathscr{U}^k, \end{cases} \tag{5.30}$$

where $\mathscr{P}^k$ and $\mathscr{U}^k$ denote $\mathscr{P}_{\mathcal{I}_\mathcal{T}(t_k),\mathcal{I}_\mathcal{V}}$ and $\times_{j\in\mathcal{I}_\mathcal{V}}\mathscr{U}_{I_k}$, respectively.

The behavior of CRH controller which constructs the state trajectory of the system, depends on the parameters of the problem and the level of uncertainties in the environment. Given the parameters introduced in the problem formulation, the time of arrivals and trajectories of targets, it is required to decide upon the value of planing horizons. In fact, the convergence of system is guaranteed only under special conditions, such as the proper choice of planing horizons.

Theorem 11. *Consider the receding horizon problem presented in* (5.30)*. Assume that* $\tilde{a} < u_{\max}$*. Then for any initial* x *in mission space and any* $\mathbf{\Pi}$ *belonging to* (5.28)*, there exists a sequence of planning horizons for the cooperative receding horizon controller where the vehicles visit targets in finite steps.*

5.4 Simulation Results

In this section, a scenario is designed to assess the performance of the proposed algorithm for an example involving two double-integrator vehicles and a set of four targets arriving at the mission space sequentially. The scenario shows the effectiveness and flexibility of the proposed method in meeting various dynamic decision criteria solely by modifying the reward functions to the most appropriate. The square $\mathcal{M} = [-200, 200] \times [-200, 200]$ is taken as the mission space which is a closed convex set. The initial position and velocity of vehicles and targets are generated randomly. Also, each of the targets have an *a priori*

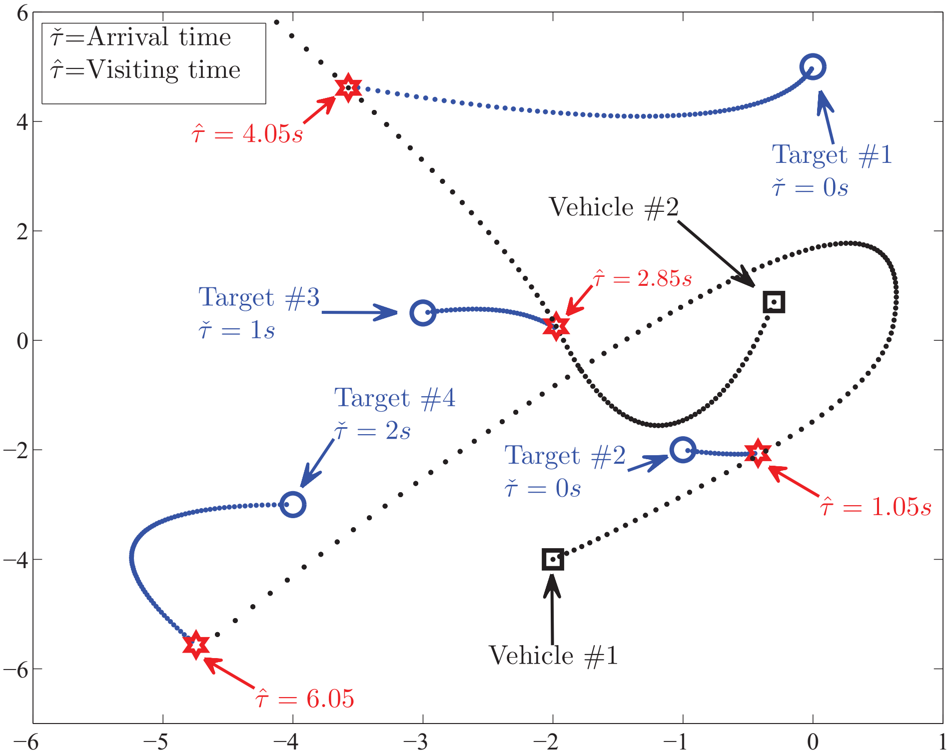

Figure 5.1: The target tracking for the vehicles and sequentially arriving targets using exponential reward function.

unknown trajectory and arrival time; however, at each time instant t the information vector of the targets $\mathbf{y}(t)$ (see (5.11)) are updated. The maximum acceleration of vehicles and targets are bounded by $u_{\max} = 2m/s$ and $\tilde{a} = 1m/s$, respectively. It is assumed that the targets always satisfy Assumptions 8 and 9. Initially two targets are present in the mission space along with the vehicles, and the remaining two targets arrive sequentially at $\{\check{\tau}_i\}_{i=3}^4 = \{2, 4\}$. The initial reward of each target is the same and equal to $\{\mathcal{R}_i = 1\}_{i=1}^4$. The targets have the same reward function which are in the form of equation (5.14) with reward discount rate parameter $\{\gamma_i = 1\}_{i=1}^4$. Fig. 5.1 depicts the position of vehicles and targets when the proposed algorithm is initialized with the above mentioned parameters and simulated until no targets remained in the mission space (Sampling time: $T_s = 0.05$). The result of this example shows that all of the targets are visited in finite time as follows: $\{\hat{\tau}_i\}_{i=1}^4 = \{1.05, 2.85, 4.05, 6.05\}$. Moreover, a total reward of $\mathcal{R} = 0.568$ is collected in this mission. Note that the assigned targets of vehicles 1 and 2 are changed to the best when new targets appears in the mission space at time moments $t = 1$ and $t = 2$.

With the help of the aforementioned simulation study, one can see the merits and efficiency of CRH controller with expected reward maximization scheme in generating the optimal assignment $\mathbf{\Pi}^k$ and control input $\mathbf{u}^k$.

Remark 14. *From Fig. 5.1, it might seem from the positions of vehicles* 1 *and* 2 *that target* 3 *should be assigned to vehicle* 1 *and target* 4 *should be assigned to vehicle* 2*. However, this is not the case because the assignment strategy takes the movement dynamics of the vehicles and targets into consideration. More precisely, not only does the strategy depend on the positions of targets and vehicles, it also depends on their velocity and acceleration vectors.*

Chapter 6

Maximum Reward Collection Problem : A Cooperative Receding Horizon Approach for Dynamic Clustering

In this chapter, the Maximum Reward Collection Problem (MRCP) in uncertain environments is investigated where multiple agents cooperate to maximize the total reward collected from a set of moving targets in the mission space with *a priori* unknown arrival times, trajectories and dynamics. The reward with respect to each of the targets has a time discounting value and can be collected only if a cluster of agents with proper number of elements visits the targets.Meanwhile, in each cluster, it is assumed that agents are able to extract a larger fraction of reward when their configuration in the cluster is close to specific configuration around the respective target. The inherited uncertainty in the environment and the dynamic clustering factor render the one-shot optimization in MRCP rather impractical. Therefore, a Cooperative Receding Horizon (CRH) controller is utilized toward maximizing the collected reward and based on the prediction of the future positions of targets with the given limited information. Some analytical aspects of problem is discussed and the effectiveness and advantages of the proposed algorithm is demonstrated via numerical simulations.

In section 6.1 the MRCP is formulated and in section 6.2 an optimization overview of the MRCP is presented. The proposed controller is introduced and formulated in section 6.3. In section 6.4 an illustrative simulation is presented.

6.0.1 Notations

Throughout the chapter, $\mathbb{N}, \mathbb{R}, \mathbb{R}_{\geq 0}$ respectively denote the set of natural numbers, real numbers, and non-negative real numbers. Also, the set of natural numbers less than or equal to n is denoted by $\mathbb{N}_n$. For a given set A and its subset B, the indicator function of B is denoted by $\mathbf{1}_B$ is a function from A to $\{0, 1\}$ and is one over B and zero elsewhere. The d dimensional Euclidean space is denoted by $\mathbb{R}^d$. Also, all-zero and all-one vectors in $\mathbb{R}^d$ are respectively represented by $\mathbf{0}_d$ and $\mathbf{1}_d$. For any vectors a and b in $\mathbb{R}^d$ inequality $\mathrm{a} \geq \mathrm{b}$ indicates that all entries of $\mathrm{a} - \mathrm{b}$ are non-negative. For any point $\mathrm{x} \in \mathbb{R}^d$ and any scalar $r \in \mathbb{R}_{\geq 0}$, the sphere with radius r centered at x is denoted by $\mathscr{B}(\mathrm{x}, r)$, and is defined as

$$\mathscr{B}(\mathrm{x}, r) = \{\mathrm{y} \in \mathbb{R}^d \mid \|\mathrm{x} - \mathrm{y}\| \leq r\}. \tag{6.1}$$

6.1 Problem Formulation

Let the *mission space* be a closed convex subset of $\mathbb{R}^d$, denoted by $\mathcal{M}$. Consider a finite number of dynamic agents, also known as vehicles, inside $\mathcal{M}$ with indices from the set $\mathcal{I}_{\mathcal{V}} = \mathbb{N}_{|\mathcal{I}_{\mathcal{V}}|}$. For any $j \in \mathcal{I}_{\mathcal{V}}$, let the dynamics of vehicle j be described by $\dot{\mathrm{x}}_j = \mathrm{u}_j(t)$, where $\mathrm{x}_j(t)$ is the position of vehicle j in the mission space at given time t and u_j is the control input for the vehicle which belongs to the set of *admissible controls*, denoted by $\mathscr{U}_{u_{\max}}$ and defined as the set of continuous functions, like $\mathrm{u} : \mathbb{R}_{\geq 0} \to \mathbb{R}^d$, bounded by $u_{\max}$ and with bounded piecewise continuous derivative.

Along with the vehicles, there exist finite number of targets with indices from $\mathcal{I}_{\mathcal{T}} = \mathbb{N}_{|\mathcal{I}_{\mathcal{T}}|}$ which arrive the mission space sequentially and move inside $\mathcal{M}$. For any $i \in \mathcal{I}_{\mathcal{T}}$, let the target i at mission space in the *a priori unknown point* $\mathrm{y}_i \in \mathcal{M}$ at the *a priori unknown time instant* $\check{\tau}_i \in \mathbb{R}_{\geq 0}$ and move afterward inside the mission space on the *a priori unknown trajectory* $\mathrm{y}_i(t)$. Without loss of generality, one can assume that the targets are indexed with respect to their arrival order, i.e. $0 \leq \check{\tau}_1 \leq \check{\tau}_2 \leq \cdots \leq \check{\tau}_{|\mathcal{I}_{\mathcal{T}}|}$.

Also, in the case that there exist initially $n_0 \in \mathbb{N}$ targets in the mission space, one has $\check{\tau}_1 = \cdots = \check{\tau}_{n_0} = 0$. Accordingly, for any $t \in \mathbb{R}_{\geq 0}$, one may define the set of indices of targets arrived up to time t, denoted by $\check{\mathcal{I}}_{\mathcal{T}}(t)$, as $\check{\mathcal{I}}_{\mathcal{T}}(t) := \{i \in \mathcal{I}_{\mathcal{T}}\ ;\ \check{\tau}_i \leq t\}$.

Each of the existing targets can be visited by a vehicle when their mutual distance is almost equal to a predetermined real scalar defined as the *visiting radius* of the target. In other words, for any $i \in \mathcal{I}_{\mathcal{T}}$ and any $j \in \mathcal{I}_{\mathcal{V}}$, vehicle j can visit the target i at time t if $r_i - \delta_{r_i} \leq \|\mathrm{x}_j(t) - \mathrm{y}_i(t)\| \leq r_i + \delta_{r_i}$ where $r_i > 0$ denotes the visiting radius of the target i and $\delta_{r_i} \in (0, r_i)$ is the *radius tolerance factor*. Also, assume that with respect to the target i there exists a time-dependent reward which can be collected if the target i is visited by a *cluster* composed of m_i number of vehicles. Here, $m_i \in \mathbb{N}$, the size of proper cluster, is the predetermined number of vehicles required for collecting the reward. Let the reward of target i be defined as the function $\mathcal{R}_i \rho_i(t)$ where $\mathcal{R}_i$ is the *initial maximum reward* and $\rho_i : \mathbb{R}_{\geq 0} \to [0, 1]$ is the *time discount function*, which is a decreasing function capturing the rate of reward loss over time. By using appropriate time discount functions, one can model different aspects of timing and scheduling such as deadlines and priorities for reward collection. As a simple example, one can consider the following function

$$\rho_i(t) = e^{-\gamma_i t}, \quad \forall\ i \in \mathcal{I}_{\mathcal{T}}, \tag{6.2}$$

for the case that there is no final hard deadline for reward collection of target i where here $\gamma_i \in \mathbb{R}_{>0}$ is the *reward discount rate parameter* for the target i. For the case that the hard deadline t_f^i is imposed on collecting the reward of target i one can take the function

$$\rho_i(t) = \max\{1 - \frac{t}{t_f^i}, 0\}, \quad \forall\, i \in \mathcal{I}_{\mathcal{T}}. \tag{6.3}$$

In this case the reward becomes zero when the deadline is passed.

With respect to each of the targets, a task is defined as collecting the respective reward which can be accomplished only by a proper cluster of vehicles. For any $i \in \mathcal{I}_{\mathcal{T}}$, one may define i^{th} *tasks accomplishment time*, denoted by $\hat{\tau}_i \in \bar{\mathbb{R}}_{\geq 0}$, as the time instant that the task i is accomplished. Note that $\hat{\tau}_i = \infty$ happens in the situations where no cluster of m_i vehicles visit target i throughout the mission time. Based on the definition of tasks accomplishment times, for any $t \in \mathbb{R}_{\geq 0}$, one can define the set of indices of *accomplished tasks* up to time t as following

$$\hat{\mathcal{I}}_{\mathcal{T}}(t) = \{i \in \mathcal{I}_{\mathcal{T}} \mid \hat{\tau}_i \leq t\}. \tag{6.4}$$

By abuse of notation, one can denote $\check{\mathcal{I}}_{\mathcal{T}}(t)$ as the set of indices of *initiated tasks* up to time t. Also, denote $\mathcal{I}_{\mathcal{T}}(t)$ as the set of indices of *current tasks* as following

$$\mathcal{I}_{\mathcal{T}}(t) = \check{\mathcal{I}}_{\mathcal{T}}(t) \backslash \hat{\mathcal{I}}_{\mathcal{T}}(t) = \{i \in \mathcal{I}_{\mathcal{T}}\ ;\ \check{\tau}_i \leq t < \hat{\tau}_i\}. \tag{6.5}$$

Regarding the trajectories of targets, note that for any $i \in \mathcal{I}_{\mathcal{T}}$, the trajectory of i^{th}

target, $\mathrm{y}_i : [\check{\tau}_i, \hat{\tau}_i] \to \mathbb{R}^d$, is assumed to be a continuously differentiable function satisfying geometric properties introduced in the sequel.

Assumption 11. *(Global Geometric Condition) For any $i \in \mathcal{I}_\mathcal{T}$ and any $\tau \in [\check{\tau}_i, \hat{\tau}_i]$, one has $\mathrm{y}_i(\tau) \in \mathcal{M}$.*

The global geometric condition ensures that once a target arrives in the mission space, it will remain inside it. This property depends not only on trajectories of targets, but also on the geometry of the mission space. For the particular case where mission space is the d-dimensional Euclidean space, the global geometric condition is immediately satisfied.

Assumption 12. *(Local Geometric Condition) There exist non-negative scalars $\tilde{v}, \tilde{a}$, such that for any $i \in \mathcal{I}_\mathcal{T}$ and any $\tau \in [\check{\tau}_i, \hat{\tau}_i]$, one has*

$$\left\| \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{y}_i(\tau) \right\| \leq \tilde{v}, \qquad \sup_{s \in (\tau, \hat{\tau}_i]} \|\alpha_i(s, \tau)\| \leq \tilde{a}, \tag{6.6}$$

where $\alpha_i(t, \tau)$ is the continuously differentiable function that for any $\tau \in [\check{\tau}_i, \hat{\tau}_i]$ and any $t \in [\tau, \hat{\tau}_i]$ the following equality holds:

$$\mathrm{y}_i(t) = \mathrm{y}_i(\tau) + \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{y}_i(\tau)(t - \tau) + \frac{1}{2} \alpha_i(t, \tau)(t - \tau)^2. \tag{6.7}$$

If function $\mathrm{y}_i(\cdot)$ is twice continuously differentiable and there exist non-negative

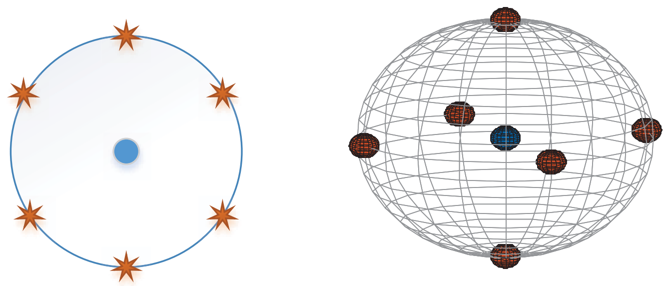

Figure 6.1: Uniform configuration of six vehicles around a target in two and three dimensional space.

scalars $\tilde{v}, \tilde{a}$ such that for any $\tau \in [\check{\tau}_i, \check{\tau}_i]$ one has

$$\|\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(\tau)\| \leq \tilde{v}, \quad \|\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{y}_i(\tau)\| \leq \tilde{a}, \tag{6.8}$$

then it can be seen from Taylor's theorem with mean-value form of the remainder [127], that $\mathrm{y}_i(t)$ satisfies Assumption 12.

Assumption 13. *For any $\tau \in \mathbb{R}_{\geq 0}$ and $i \in \mathcal{I}_{\mathcal{T}}(\tau)$, the position and velocity vectors of target i are given at the beginning of each time horizon, i.e. at time instant τ in* (6.7).

Remark 15. *For any $\tau \in \mathbb{R}_{\geq 0}$ and $i \in \mathcal{I}_{\mathcal{T}}(\tau)$, using the Assumption 13, one can estimate the positions of the target i for any future instant $t \in [\tau, \hat{\tau}_i]$ as $\hat{\mathrm{y}}_i(t) = \mathrm{y}_i(\tau) + (t - \tau)\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{y}_i(\tau)$.*

The vehicles collect some amount of reward by accomplishing each of the tasks.

It is assumed that the amount of each of these collected rewards depends on the configuration of the vehicles around the target and the time instant. In other words, the vehicles collecting the reward of a target can collect larger fraction of that when their configuration is closer to the uniform distribution over the sphere centered at the target and with radius equal to the visiting radius (Figure 6.1). More precisely, for any $m \in \mathcal{I}_{\mathcal{T}}$, let $\mathfrak{f}_m : \times_{j=1}^{m} \mathbb{R}^d \to [0, 1]$ be a function such that for any $\mathrm{z}_1, \ldots, \mathrm{z}_m \in \mathbb{R}^d$. The value of $\mathfrak{f}_m(\mathrm{z}_1, \ldots, \mathrm{z}_m)$ shows the proximity of distribution of the points $\mathrm{z}_1, \ldots, \mathrm{z}_m$ to uniform distribution of m points on the unit sphere in $\mathbb{R}^d$ and also, it becomes equal to one and takes the maximum when the distribution of the points $\mathrm{z}_1, \ldots, \mathrm{z}_m$ be exactly as uniform distribution of m points on the unit sphere in $\mathbb{R}^d$. Accordingly, for any $i \in \mathcal{I}_{\mathcal{T}}$, one can define the function $\mathfrak{f}_i$ for the fraction of reward collected at the i^{th} task accomplishment time as following

$$\mathfrak{f}_i(\mathbf{x}(\hat{\tau}_i; \mathbf{x}_0, \mathrm{u})) = \max_{\mathcal{J} \subseteq \mathcal{I}_{\mathcal{V}}, |\mathcal{J}| = m_i} \mathfrak{f}_m(\frac{1}{r_i}\big(\mathrm{x}_j(\hat{\tau}_i) - \mathrm{y}_i(\hat{\tau}_i)\big)_{j \in \mathcal{J}}), \tag{6.9}$$

where $\mathbf{x}(\cdot; \mathbf{x}_0, \mathrm{u})$ is the solution of total system, starting from $\mathbf{x}_0$ and applying the control input u.

Now, one can define the *mission* as the procedure of controlling the vehicles for cooperatively collecting the maximum possible rewards from the targets. Considering the uncertainties and the limitations on information in the introduced paradigm, it is desired to obtain a near-optimal reward collection cooperative control policy for mission accomplishment, which is discussed in subsequent sections.

6.2 An Optimization Overview

Let the *total reward function*, denoted by $\mathfrak{R}_\Sigma$, be a function like $\mathfrak{R}_\Sigma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ where for any $t \in \mathbb{R}_{\geq 0}$, the $\mathfrak{R}_\Sigma(t)$ accounts for the net reward available at the time t and defined as following

$$\mathfrak{R}_\Sigma(t, \mathrm{u}) := \sum_{i \in \mathcal{I}_\mathcal{T}} \mathscr{R}_i \rho_i(t) \mathbf{1}_{[\check{\tau}_i, \hat{\tau}_i)}(t). \tag{6.10}$$

The dependency of the total reward function on u is through the $(\hat{\tau}_i)_{i \in \mathcal{I}_\mathcal{T}}$ which itself depends on trajectories of targets and also trajectories of vehicles that are subsequently dependent on u. Considering the initial conditions and the uncertainties in the problem, one should note that collecting all of the total rewad may not be feasible. To maximize the total reward, one can define the *total collected reward* as a function of control u as

$$\mathfrak{R}_\infty(\mathrm{u}) := \sum_{i \in \mathcal{I}_\mathcal{T}} \mathscr{R}_i \rho_i(\hat{\tau}_i) f_i(\mathbf{x}(\hat{\tau}_i; \mathbf{x}_0, \mathrm{u})). \tag{6.11}$$

where the function $f_i(\cdot)$ measures the closeness of the vehicles distribution to the uniform distribution of m_i points over a sphere centered at $\mathbf{y}_i(\hat{\tau})$ with radius r_i, for any $i \in \mathcal{I}_\mathcal{T}$. Equation (6.11) shows that the maximum reward collection problem can be formulated as an optimization problem defined as $\max_{\mathrm{u} \in \mathfrak{U}} \; \mathfrak{R}_\infty(\mathrm{u})$, where $\mathfrak{U}$ denotes the set of admissible control inputs for all the vehicles, i.e. $\mathfrak{U} = \bigtimes_{j \in \mathcal{I}_\mathcal{V}} \mathscr{U}_{u^j_{\max}}$. Note that $(\hat{\tau}_i)_{i \in \mathcal{I}_\mathcal{T}}$ not only depends on $\mathbf{x}$ and subsequently on u, but also on $\mathbf{y}(\cdot)$. This dependency renders the optimization problem intractable since the feasible set is infinite-dimensional and there exists uncertainty in the problem. Therefore, it is preferred to obtain a less computationally demanding alternative to the introduced optimization problem, such as the time decomposition based method of receding horizon scheme. Subsequently, it is essential to design an alternative payoff function accounting for estimation of total reward and also clustering strategy and uniform configuration in vicinity of the targets. The design of relative receding horizon scheme, the payoff function and the respective appropriate feasible sets are presented in the next section.

6.3 Cooperative Receding Horizon Scheme

In this section, a proper cooperative receding horizon (CRH) controller is developed to generate the paths for the vehicles and obtain desired configurations. The controller generates headings and step sizes for the vehicles iteratively. At each time instants, $\{t_k\}_{k\in\mathcal{K}} \in \mathbb{R}_{\geq 0}$, the information relative to targets and the vehicles is updated and also, an optimization problem is formulated with a payoff function which estimates the collected reward by the end of mission and assesses deviation of clustering and configuration of vehicles from proper ones. Finally, the desired control input at time instant t_k, denoted by $\mathbf{u}^k = (\mathrm{u}_j(t_k))_{j\in\mathcal{I}_\mathcal{V}}$, is provided from solution of the optimization problem, for any $k \in \mathcal{K}$.

6.3.1 Reward Prediction

Denote H_k as the *planing horizon* in k^{th} of CRH controller, i.e. $H_k := t_{k+1} - t_k$, for any $k \in \mathcal{K}$. Let the control input $\mathbf{u}^k = (\mathrm{u}_j(t_k))_{j\in\mathcal{I}_\mathcal{V}}$ be applied to the vehicles, in the time interval $[t_k, t_k + H_k)$. Then, for any $H \in [0, H_k]$, it follows from dynamics of vehicles that the positions of vehicles at the time $t_k + H$ are given by

$$\mathbf{x}(t_k + H) = \mathbf{x}(t_k) + \mathrm{u}(t_k)H. \tag{6.12}$$

Also, from Remark 15 and the available information at time instant t_k, one can estimated the positions of targets at time $t_k + H$, as following

$$\hat{\mathbf{y}}(t_k + H) = \hat{\mathbf{y}}(t_k) + H\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{y}}(t_k). \tag{6.13}$$

Remark 16. *One might note that the error of the estimation given in* (6.13) *is bounded by* $\frac{1}{2}H_k^2\tilde{\boldsymbol{a}}$ *for each entry of* $\mathbf{y}(t_k+H)$. *Hence, for desired estimation accuracy, it is enough to set* H_k *small enough. More precisely, being* $\frac{1}{2}H_k^2\tilde{\boldsymbol{a}}$ *comparatively smaller than* $H_k\tilde{v}$*, or equivalently* $H_k \ll 2\tilde{v}\tilde{\boldsymbol{a}}^{-1}$*, one can disregard the estimation error term* $\frac{1}{2}H_k^2\tilde{\boldsymbol{a}}$*.*

Let $\tau \in \mathbb{R}_{\geq 0}$ be a time instant such that $\tau \geq t_{k+1}$. Considering the uncertainties in the trajectories of targets, and also the fact that the available information on them is very limited, for any $i \in \mathcal{I}_\mathcal{T}$, one may model $\mathrm{y}_i(\tau)$ as sum of $\mathrm{y}_i(t_{k+1})$, the position of targets i at time t_{k+1}, and a random vector, denoted by $\mathscr{E}_i^{k+1}(\tau)$, with radially symmetric distribution and values in $\mathscr{B}(\mathbf{0}_d, (\tau - t_{k+1})v)$. From properties of $\mathscr{E}_i^{k+1}(\tau)$, it follows that $\mathbb{E}[\mathscr{E}_i^{k+1}(\tau)] = \mathbf{0}_d$. Based on this, one can best estimate $\mathrm{y}_i(\tau)$ by $\mathrm{y}_i(t_{k+1})$.

For any $i \in \mathcal{I}_\mathcal{T}$ and $j \in \mathcal{I}_\mathcal{V}$, define the *expected reaching time* of vehicle j to target i, denoted by $\tau_{ij}(\mathrm{u}^k, t_k)$, as the estimation of the time that vehicle j is expected to reach the target i, assuming that the control inputs u^k is applied at time t_k for a planned horizon H_k and, from the time instant t_{k+1}, the vehicle j takes the responsibility of the target i and be assigned to it. From the given estimations, it follows that

$$\tau_{ij}(\mathrm{u}^k, t_k) = t_k + H_k + \frac{\|\mathrm{x}_j(t_k + H_k) - \hat{\mathrm{y}}_i(t_k + H_k)\| - r_i}{V_j}, \tag{6.14}$$

when $\|\mathrm{x}_j(t_k + H_k) - \hat{\mathrm{y}}_i(t_k + H_k)\| > r_i$, and otherwise $\tau_{ij}(\mathrm{u}^k, t_k) = t_k + H_k$. Subsequently, the vehicle j expects to collect the respective reward from the target i, estimated as $\mathscr{R}_i\tilde{\rho}_{ij}(\mathrm{u}^k, t_k)$ where $\tilde{\rho}_{ij}(\mathrm{u}^k, t_k)$ is defined as

$$\tilde{\rho}_{ij}(\mathrm{u}^k, t_k) := \rho_i[\tau_{ij}(\mathrm{u}^k, t_k)]. \tag{6.15}$$

This gives the reward prediction for a pair of vehicle and target. Reward prediction for the whole team is discussed next.

6.3.2 Clustering and Task Assignments

In each iteration of CRH controller, vehicles are expected to decide on their clusterings and the task assignment strategies, based on the available information, and subsequently, plan their paths. In order to characterize this decision-makings procedure, the notions of clustering strategy and task assignment are discussed below.

Clustering Strategies

For any $i \in \mathcal{I}_{\mathcal{T}}$, the *clustering strategy factor* for the target i is characterized as a real scalar in $[0, 1]$, denoted by c_i, which reflects the level of responsibility of vehicles for configuring a proper clustering around the target i in order to collect its reward. Note that at any point in time, each of the clustering strategy factors depends on the positions of all the vehicles and targets. More precisely, for any $i \in \mathcal{I}_{\mathcal{T}}(t)$, the clustering strategy c_i can be represented as a function of the following form:

$$c_i : \mathcal{M}^{|\mathcal{I}_{\mathcal{V}}|} \times \mathcal{M}^{|\mathcal{I}_{\mathcal{T}}(t)|} \to [0, 1]. \tag{6.16}$$

Note that the vehicles should not accept responsibilities more than they can handle. More specifically, for any $k \in \mathcal{K}$, one must have that

$$\sum_{i \in \mathcal{I}_{\mathcal{T}}(t_k)} m_i c_i(\mathbf{x}^k, \mathbf{y}^k) \leq m, \tag{6.17}$$

where m_i is the required number of vehicles for the cluster respective to target i, for any $i \in \mathcal{I}_{\mathcal{T}}(t_k)$, and $\mathbf{x}^k, \mathbf{y}^k$ are the positions of vehicles and positions of targets at t_k, respectively. For convenience in the notations, denote c_i^k as $c_i(\mathbf{x}^k, \mathbf{y}^k)$ and $\mathbf{c}^k$ as the vector $(c_i^k)_{i \in \mathcal{I}_{\mathcal{T}}(t_k)}$.

Task Assignments

Similar to the clustering strategy factors, for any $i \in \mathcal{I}_{\mathcal{T}}$ and $j \in \mathcal{I}_{\mathcal{V}}$, an *assignment* of task i to vehicle j is defined as a real scalar in $[0, 1]$, denoted by a_{ij} which shows the amount of interest of vehicle j in being task i assigned to it and depends on the positions of vehicles and the positions of targets. Specifically, for any $i \in \mathcal{I}_{\mathcal{T}}(t)$, the assignment a_{ij} is the following function:

$$a_{ij} : \mathcal{M}^{|\mathcal{I}_{\mathcal{V}}|} \times \mathcal{M}^{|\mathcal{I}_{\mathcal{T}}(t)|} \rightarrow [0, 1]. \tag{6.18}$$

Various methods such as Voronoi-based assignment policy [134] and competition-based assignment [108] can be exploited to design the assignment functions. The assignments are required to have some desired structures. First, each of the vehicles is expected to consider all of the current tasks and also, the option of being assigned to none of the tasks. Subsequently, the sum of task assignments for each vehicle must be less than or equal one, i.e.

$$\sum_{i\in\mathcal{I}_{\mathcal{T}}(t_k)} a_{ij}(\mathbf{x}^k,\mathbf{y}^k) \leq 1, \quad \forall j \in \mathcal{I}_{\mathcal{V}}, \tag{6.19}$$

where the sum is zero if $\mathcal{I}_{\mathcal{T}}(t_k) = \emptyset$. Also, the vehicles are expected to adapt their assignments to each of the targets in accordance to the responsibility accepted for the target and also, the required number of vehicles to collect its reward. More precisely, the assignments and the clustering strategy factor are required to satisfy the following identity:

$$\sum_{j\in\mathcal{I}_{\mathcal{V}}} a_{ij}(\mathbf{x}^k,\mathbf{y}^k) = m_i c_i(\mathbf{x}^k,\mathbf{y}^k), \quad \forall i \in \mathcal{I}_{\mathcal{T}}(t_k). \tag{6.20}$$

For convenience of notation, $a_{ij}(\mathbf{x}^k,\mathbf{y}^k)$ and the matrix $\mathbf{A}(\mathbf{x}^k,\mathbf{y}^k)$ having $a_{ij}(\mathbf{x}^k,\mathbf{y}^k)$ as its entry at i^{th} row and j^{th} column, for any $i \in \mathcal{I}_{\mathcal{T}}(t_k)$ and $j \in \mathcal{I}_{\mathcal{V}}$, are shown by a_{ij}^k and $\mathbf{A}^k$.

6.3.3 Potential Function

Here, the design of potential function for the optimization problem in the receding horizon scheme is introduced. The potential function consists of two main parts, one for the maximum total reward expected to be collected by the end of mission, and one for proper configurations of vehicles.

Total Expected Reward

Let $i \in \mathcal{I}_{\mathcal{T}}$ and $j \in \mathcal{I}_{\mathcal{V}}$ and also the control inputs u^k be applied at time t_k for a planned horizon H_k. Similar to discussion given in section 6.3.1, one can estimate the $a_{ij}(\mathbf{x}(\hat{\tau}_{ij}(\mathbf{u}^k, t_k)), \hat{\mathbf{y}}(\hat{\tau}_{ij}(\mathbf{u}^k, t_k)))$ by $a_{ij}(\mathbf{x}(t_{k+1}), \hat{\mathbf{y}}(t_{k+1}))$. Define the function $\tilde{a}_{ij}(\mathbf{u}^k, t_k)$ as following:

$$\tilde{a}_{ij}(\mathbf{u}^k, t_k) = a_{ij}(\mathbf{x}(\hat{\tau}_{ij}(\mathbf{u}^k, t_k)), \hat{\mathbf{y}}(\hat{\tau}_{ij}(\mathbf{u}^k, t_k))). \tag{6.21}$$

Define $\mathfrak{R}^{k+1}$ as the maximum total reward which the vehicles expects at time t_{k+1} to collect by the end of the mission. Considering the equations (6.15), (6.21), the assignments and the clustering strategy factors, one can estimate $\mathfrak{R}^{k+1}$ as

$$\mathfrak{R}^{k+1}(\mathbf{u}^k, t_k) = \sum_{j \in \mathcal{I}_{\mathcal{V}}} \sum_{i \in \mathcal{I}_{\mathcal{T},j}(t_k)} \frac{\mathcal{R}_i}{m_i} \tilde{c}_i(\mathrm{u}^k, t_k) \tilde{\rho}_{ij}(\mathrm{u}^k, t_k) \tilde{a}_{ij}(\mathbf{u}^k, t_k), \tag{6.22}$$

where $\tilde{c}_i(\mathrm{u}^k, t_k)$ is defined similar to $\tilde{\rho}_{ij}(\mathrm{u}^k, t_k)$ and $\tilde{a}_{ij}(\mathbf{u}^k, t_k)$, i.e.

$$\tilde{c}_i(\mathrm{u}^k, t_k) = \tilde{c}_i(\mathbf{x}(\hat{\tau}_{ij}(\mathbf{u}^k, t_k)), \hat{\mathbf{y}}(\hat{\tau}_{ij}(\mathbf{u}^k, t_k))) \tag{6.23}$$

Proper Configurations

The main objective, when the vehicles are distant from their intended targets, is the interception, and when they are in the vicinity of them, is to obtain the proper configuration. Therefore, we define the *mollifier* distance function, denoted by ϕ, as

$$\phi(x) = \begin{cases} e^{-\frac{x^2}{1-x^2}}, & \text{if } |x| < 1, \\ 0, & \text{if } |x| \geq 1, \end{cases} \tag{6.24}$$

and the *smooth step function*, denoted by ψ, as

$$\psi(x) = \frac{\chi(x)}{\chi(x) + \chi(1-x)}, \tag{6.25}$$

where the function χ is defined as following

$$\chi(x) = \begin{cases} e^{-\frac{1}{x}}, & \text{if } x > 0, \\ 0, & \text{if } x \leq 0. \end{cases} \tag{6.26}$$

These functions are shown in Figure 6.2. Note that the functions ϕ, ψ and χ are non-analytic infinite-time differentiable functions. Based on these functions, for any $i \in \mathcal{I}_\mathcal{T}$

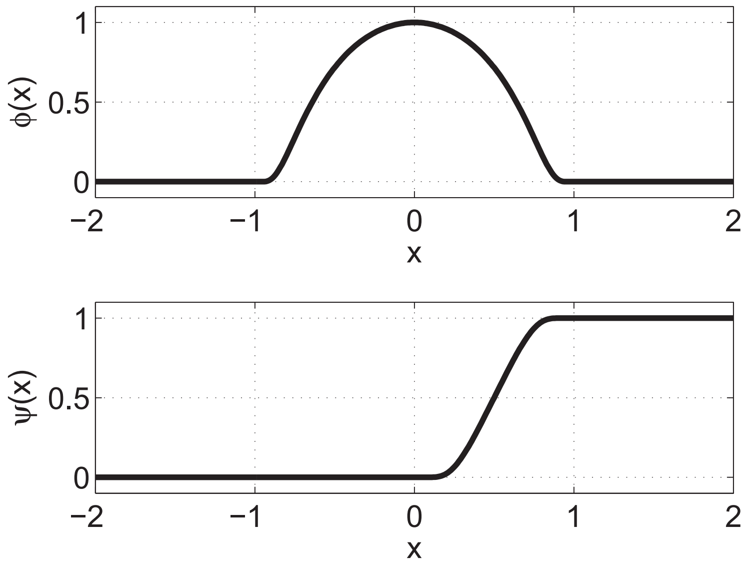

Figure 6.2: The mollifier and smooth step function.

and $\delta_{r_i} \in \mathbb{R}_{>0}$, one can define the function $\phi_i : \mathbb{R}^d \to \mathbb{R}$ as

$$\phi_i(\mathrm{x}, \delta_{r_i}) = \phi(\frac{\|\mathrm{x} - \mathrm{y}_i\| - r_i}{\delta_{r_i}}). \tag{6.27}$$

The value of $\phi_i(\mathrm{x}, \delta_{r_i})$ is always non-negative and non-zero only if $\|\mathrm{y}_i - x\| \in (r_i - \delta_{r_i}, r_i + \delta_{r_i})$. Similarly to ϕ_i, the function ψ_i can be defined, for any $i \in \mathcal{I}_{\mathcal{T}}$ and $\delta_{r_i} \in \mathbb{R}_{>0}$, as following

$$\psi_i(\mathrm{x}, \delta_{r_i}) = 1 - \psi(\frac{\|\mathrm{x} - \mathrm{y}_i\| - r_i}{\delta_{r_i}}). \tag{6.28}$$

Note that for any $\mathrm{x} \in \mathbb{R}^d$, one has $\psi_i(\mathrm{x}, \delta_r) \in [0, 1]$, and also, $\psi_i(\mathrm{x}, \delta_r) = 0$ if and only if $\|\mathrm{y}_i - x\| \geq r_i + \delta_{r_i}$ and $\psi_i(\mathrm{x}, \delta_{r_i}) = 1$ if and only if $\|\mathrm{y}_i - x\| \leq r_i$.

For any $i \in \mathcal{I}_{\mathcal{T}}$, let $\mathfrak{d}_i$ be a function defined as

$$\mathfrak{d}_i(\mathrm{x}) = \left(1 + (\|\mathrm{x} - \mathrm{y}_i\| - r_i)^2\right)^{-1}.$$

It can be easily verified that $\mathfrak{d}_i$ takes its maximum value only if $\|\mathrm{x} - \mathrm{y}_i\| = r_i$, i.e. when x has the desired distance from y_i. A proper potential function can be obtained in order to force the vehicles to take the desired distances from the targets during the reward collection. Consider the function $\tilde{\mathfrak{d}}_i$ as $\tilde{\mathfrak{d}}_i(\mathrm{u}^k, t_k) = \mathfrak{d}_i(\mathrm{x}_j^k + H_k \mathrm{u}_j^k)$, where u^k is the control inputs applied at time t_k for a planned horizon H_k. The desired potential function is defined as

$$\mathfrak{D}^{k+1}(\mathrm{u}^k, t_k) = \sum_{i \in \mathcal{I}_{\mathcal{T}}(t_k)} \sum_{j \in \mathcal{I}_{\mathcal{V}}} \tilde{c}_i(\mathrm{u}^k, t_k) \tilde{a}_{ij}(\mathrm{u}^k, t_k) \tilde{\mathfrak{d}}_i(\mathrm{u}^k, t_k) \tilde{\psi}_{ij}(\mathrm{u}^k, t_k), \tag{6.29}$$

where $\tilde{\psi}_i$ is defined as $\tilde{\psi}_{ij}(\mathrm{u}^k, t_k) = \psi_i(\mathrm{x}_j^k + H_k \mathrm{u}_j^k, \delta_{r_i})$ for a given $\delta_{r_i} \in \mathbb{R}_{>0}$.

The vehicles, besides taking the desired distances from their intended targets, are supposed to configure so as to have an almost uniform distribution over the sphere centred at respective the intended targets. Note that the distribution of a set of points on a sphere is uniform distribution when the sum of their mutual distances is maximum. One can obtain an appropriate potential function forcing the vehicles for the configurations having the uniform distribution. Given that the control inputs u^k are applied at time t_k

for a planned horizon H_k, one can define the proper potential function as following

$$\mathfrak{F}^{k+1}(\mathrm{u}^k, t_k) = \sum_{i\in\mathcal{I}_{\mathcal{T}}(t_k)} \sum_{j,l\in\mathcal{I}_{\mathcal{V}}} \tilde{c}_i(\mathrm{u}^k, t_k)\tilde{a}_{ij}(\mathrm{u}^k, t_k)\tilde{a}_{il}(\mathrm{u}^k, t_k)\tilde{\phi}_{ij}(\mathrm{u}^k, t_k)\tilde{\phi}_{il}(\mathrm{u}^k, t_k) \\ \|\mathrm{x}_j^k - \mathrm{x}_l^k + H_k(\mathrm{u}_j^k - \mathrm{u}_l^k)\|, \tag{6.30}$$

where $\tilde{\phi}_{ij}$ and $\tilde{\phi}_{il}$ are defined as $\tilde{\phi}_{ij}(\mathrm{u}^k, t_k) = \phi_i(\mathrm{x}_j^k + H_k\mathrm{u}_j^k, \delta_{r_i})$ and $\tilde{\phi}_{il}(\mathrm{u}^k, t_k) = \phi_i(\mathrm{x}_l^k + H_k\mathrm{u}_l^k, \delta_{r_i})$, respectively, for a given $\delta_{r_i} \in \mathbb{R}_{>0}$.

Clustering Imperfection

The vehicles are allowed to collect the reward of a target only if the respective cluster has the required number of vehicles, i.e. the respective cluster is perfect. Hence, the vehicles are supposed to establish only perfect clusters. To this end, one can define a potential function, as a cost for imperfect clustering, which considers the imperfection in each of the established clusters and the grade of responsibility accepted for their respective targets. Accordingly, one can define the *imperfection cost function* as

$$\mathfrak{I}^{k+1}(\mathrm{u}^k, t_k) = \sum_{i\in\mathcal{I}_{\mathcal{T}}(t_k)} \tilde{c}_i(\mathrm{u}^k, t_k)\Big(m_i - \sum_{j\in\mathcal{I}_{\mathcal{V}}} \tilde{a}_{ij}(\mathrm{u}^k, t_k)\Big) \tag{6.31}$$

where $\tilde{c}_i$ and $\tilde{a}_{ij}$ are defined as before and also, it is assumed that the control inputs u^k are applied at time t_k for a planned horizon H_k.

6.3.4 Cooperative Receding Horizon Trajectory Construction

The receding horizon scheme controller provides the inputs by solving an optimization problem in each iteration. The formulation of this optimization problem is discussed below.

Considering the potential functions introduced, one can define the payoff function as

$$\mathfrak{J}^{k+1}(\mathrm{u}^k, t_k) = \omega_{\mathfrak{R}}\mathfrak{R}^{k+1}(\mathrm{u}^k, t_k) + \omega_{\mathfrak{D}}\mathfrak{D}^{k+1}(\mathrm{u}^k, t_k) + \omega_{\mathfrak{F}}\mathfrak{F}^{k+1}(\mathrm{u}^k, t_k) - \omega_{\mathfrak{I}}\mathfrak{I}^{k+1}(\mathrm{u}^k, t_k), \quad (6.32)$$

where $\omega_{\mathfrak{R}}$, $\omega_{\mathfrak{D}}$, $\omega_{\mathfrak{F}}$ and $\omega_{\mathfrak{I}}$ are the non-negative real-valued weights for the respective terms. In addition, $\|\mathrm{u}_j^k\| \leq u_{\max}$, for any $j \in \mathcal{I}_{\mathcal{V}}$, i.e. u^k belongs to the set of *admissible control inputs* denoted by $\mathcal{U}^k$ and defined as

$$\mathcal{U}^k = \{\mathbf{u} = (\mathrm{u}_j)_{j\in\mathcal{I}_{\mathcal{V}}} \ ; \ \mathrm{u}_j \in \mathbb{R}^d, \|\mathrm{u}_j\| \leq u_{\max}, \forall j \in \mathcal{I}_{\mathcal{V}}\}. \quad (6.33)$$

Besides these explicit constraints, there are other implicit constraints imposed on u^k through $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{c}}$, where $\tilde{\mathbf{A}}$ is defined as

$$\tilde{\mathbf{A}}(\mathrm{u}^k, t_k) = \left(\tilde{a}_{ij}^k(\mathrm{u}^k, t_k)\right)_{i\in\mathcal{I}_{\mathcal{T}}(t_k), j\in\mathcal{I}_{\mathcal{V}}}, \quad (6.34)$$

and $\tilde{\mathbf{c}}$ is defined as

$$\tilde{\mathbf{c}}(\mathrm{u}^k, t_k) = \big(\tilde{c}_i^k(\mathrm{u}^k, t_k)\big)_{i\in\mathcal{I}_{\mathcal{T}}(t_k)}. \tag{6.35}$$

Specifically, if one defines the set $\mathscr{F}^k$ as

$$\mathscr{F}^k = \{(\mathbf{A}, \mathbf{c}) \in [0,1]^{|\mathcal{I}_{\mathcal{T}}(t_k)|\times|\mathcal{I}_{\mathcal{V}}|} \times [0,1]^{|\mathcal{I}_{\mathcal{T}}(t_k)|}; \mathbf{A}\mathbf{1}_{|\mathcal{I}_{\mathcal{V}}|} = \mathbf{1}_{|\mathcal{I}_{\mathcal{T}}(t_k)|},\ \mathbf{A}^{\mathsf{T}}\mathbf{1}_{|\mathcal{I}_{\mathcal{T}}(t_k)|} = \mathrm{m}^k\}, \tag{6.36}$$

where m^k is defined as $(m_i)_{i\in\mathcal{I}_{\mathcal{T}}(t_k)}$, then it is supposed to have

$$\big(\tilde{\mathbf{A}}(\mathrm{u}^k, t_k), \tilde{\mathbf{c}}(\mathrm{u}^k, t_k)\big) \in \mathscr{F}^k. \tag{6.37}$$

Note that the implicit dependency of $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{c}}$ on u^k is through the optimization problem.

From the given discussion, the k^{th} iteration in CRH scheme, P_k, can be written as

$$\mathrm{P}_k: \quad \begin{cases} \max & \mathfrak{J}^{k+1}(\mathbf{u}^k, t_k) \\ \text{s.t.} & \mathbf{u}^k \in \mathcal{U}^k, \\ & \big(\tilde{\mathbf{A}}(\mathrm{u}^k, t_k), \tilde{\mathbf{c}}(\mathrm{u}^k, t_k)\big) \in \mathscr{F}^k. \end{cases} \tag{6.38}$$

The control input u^k is obtained by solving the optimization P_k, given in (6.38).

6.3.5 Analysis of CRH Scheme

The trajectories of vehicles are constructed iteratively from the solutions of (6.38). Hence, the behavior of the system depends on the optimization problem presented in (6.38).

Definition 13. *The trajectory* $\mathbf{x}(t)$ *is called a stationary trajectory if for all* $i \in \mathcal{I}_{\mathcal{T}}$*, the* i^{th} *target hitting time or* i^{th} *task completion time is finite, i.e. one has* $\check{\tau}_i < \infty$*.*

Regarding the behavior of the system, the stationarity of vehicle's trajectory can be guaranteed under some assumptions and conditions given in the sequel.

Assumption 14. *There exists positive real scalars* δ_{r_i} *such that for any* $t \in \mathbb{R}_{>0}$ *and any distinct* $i, i' \in \mathcal{I}_{\mathcal{T}(t)}$*, one has*

$$\|\mathrm{y}_i(t) - \mathrm{y}_{i'}(t)\| > r_i + r_{i'} + \delta_{r_i} + \delta_{r_{i'}}. \tag{6.39}$$

Proposition 1. *Let Assumption 14 hold,* $\tilde{v} < u_{\max}$ *and positive real scalars* $\omega_{\mathfrak{R}}$*,* $\omega_{\mathfrak{D}}$*,* $\omega_{\mathfrak{F}}$ *and* δ_{r_i}*, for any* $i\mathcal{I}_{\mathcal{T}}$*, be given. Then, there exist* $\underline{\omega}_{\mathfrak{J}} \in \mathbb{R}_{>0}$ *such that for any* $\underline{\omega}_{\mathfrak{J}} \leq \omega_{\mathfrak{J}}$*, one can obtain* $K \in \mathbb{N}$ *and* $\{H_k\}_{k=1}^{K}$ *where the trajectories of vehicles constructed by the resulting CRH scheme are stationary.*

6.4 Simulation Results

In this section, the performance of the proposed method is investigated through a simulation study. The simulation scenario involves eight vehicles and a set of four targets arriving sequentially in the mission space which is a closed convex set in a flat plane $\mathcal{M} = [-200, 200] \times [-200, 200]$. The initial position of the vehicles and targets are produced randomly with uniform distribution. The arrival time of the targets is also assumed to be a random variable with exponential distribution and the rate parameter $\lambda = 1$. Targets have *a priori* unknown trajectories with the maximum velocity of $\tilde{v} = 15m/s$ and the maximum velocity of vehicles is $u_{\max} = 30m/s$. For generality, the targets' trajectories are chosen randomly. Initially, along with the vehicles, two targets are also present in the mission space, and the remaining two targets arrive sequentially at $\{\check{\tau}_i\}_{i=3}^4 = \{1.24, 2.62\}$. The number of vehicles needed to cluster around each target is $\{m_i\}_{i=1}^4 = \{5, 4, 5, 4\}$. The vehicles should both maintain a distance of $\{r_i\}_{i=1}^4 = 30$ from the targets with $\delta_r = 1.5$ and try to surround it uniformly i.e. the distance between each two neighbour vehicles in the cluster should be the same. The weights of the payoff function (6.32) are set to $\omega_{\mathfrak{R}} = \omega_{\mathfrak{D}} = 1, \omega_{\mathfrak{F}} = 10, \omega_{\mathfrak{I}} = 5000$ and the finally reward function of all of the targets are the same and in the form of (6.3) with initial reward $\{\mathcal{R}_i = 100\}_{i=1}^4$ and deadline of $\{t_f^i\}_{i=1}^4 = 50$.

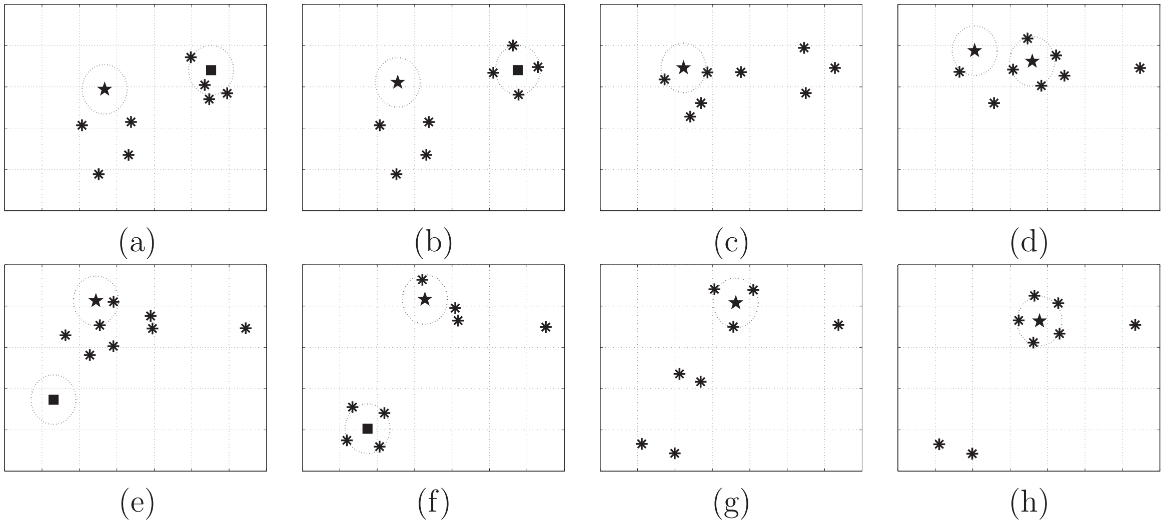

Figure 6.3: The result of cooperative recoding horizon maximum award collecting problem solved for eight vehicles marked by ⁎ and four targets. The first and second target, depicted by ★ and ■, are appeared in the mission space from the start. The third and fourth targets, represented by ■, arrived subsequently

Figure 6.3 shows snap shots of different stages of solving the maximum reward collecting problem. Only the critical decision making act of the vehicles are demonstrated in this figure. Part (a) shows the initial state of the targets and vehicles in the mission space. The transition between part (a) and (b) of this figure demonstrates the effect of $\mathfrak{I}^{k+1}$ with such a high $\omega_{\mathfrak{I}}$ coefficient in the payoff function which led to $\tilde{c}_1 = 0, \tilde{c}_2 = 1$. This happened because with the available number of vehicles at that moment, either the first or the second target could be visited and in that state of the mission space, the second target offered more reward because it was closer. After visiting the second target the vehicles aimed to the first target in part (c), however, target number three appeared in the mission space and the vehicles changed their behaviour and first surrounded target three. This flexibility in selecting a cluster to join and a target to visit happens again when the fourth target arrived in the mission space in part (e) and the vehicles abandon the first target again and visit the fourth target. As one can see, in the first target is the last one to be visited and this can be explained by the fact that rewards are decreasing in time and the vehicles are insufficient to intercept two targets.

Chapter 7

Summary and Extensions

This chapter provides a brief summary of the contributions of the thesis in Section 7.1, and then some suggestions for future research direction in this area are given in Section 7.2.

7.1 Summary of Contributions

In this thesis, a cooperative receding horizon scheme is developed, which uses a time decomposition approach to design a controller for the multi-target interception problem in an uncertain environment where each of the targets arrive in the mission space sequentially at *a priori* unknown arrival times, in *a priori* unknown positions and moving on *a priori* unknown trajectories.

In Chapter 2 of the thesis, a time decreasing reward was assigned with each target which can be collected only if the target is visited by at least one vehicle. The team objective is to maximize the total collected rewards. At each iteration, the vehicles encounter multiple targets, some of which could be new in the mission space. Each target has an *a priori* unknown trajectory with a bounded velocity. As the targets arrive in the mission space sequentially, vehicles aim at visiting them in minimum time to avoid a burst of unvisited target population and at the same time to have a stationary state. Accordingly, a cooperative receding horizon controller is designed to collect maximum possible rewards, and hence, to track moving targets with *a priori* unknown dynamics using a team of vehicles by maximizing the expectation of total collectible rewards.

In the Chapter 3, the paradigm introduced above is extended to a receding-horizon-based dynamic decision-making controller for control of a single vehicle toward intercepting a group of infinite number of targets which were arrive in the mission space sequentially in *a priori* unknown locations. They then move with unspecified trajectories and unknown dynamics. The arrival times of the targets are modeled stochastically by a renewal process. Convergence analysis is provided, and simulations results are given for different scenarios, e.g., frequent and infrequent target arrivals.

Then, the cooperative receding horizon controller designed in the first part is extended in Chapter 4 to the case where vehicles have limited ranges for sensing the targets, and also limited ranges for communication. This is accomplished using a game-theoretic approach. In this method, a utility function is designed for each vehicle, which depends on the rewards as well as the vehicles' constraints. The resulting structure forms a potential game, where the the total collectible reward is the potential function. Using appropriate learning dynamics, vehicles decide upon their strategies and move in proper directions accordingly.

In Chapter 5, using some important concepts from optimal control theory, the reward assignment strategy is extended to double-integrator vehicles. At each iteration, a time optimal control problem is considered for each pair of vehicles and targets, and then solved by Pontryagin's maximum principle. Using the solution of these optimal control problems, an estimation of the total collectible reward is obtained and introduced as the payoff function for reward maximization. It is shown that control inputs obtained from the solution of the resulting optimization problems generate stationary trajectories.

In Chapter 6, the cooperative receding horizon scheme introduced in Chapter 2 is extended to case where agents are dynamically clustered and assigned to the targets to collect rewards. The introduced payoff functions account for the estimation of maximum total reward expected to be collected by the end of mission, the clustering and assignment strategies, uniform configurations of agents in the vicinity of the targets, and finally, how imperfect the cluster are.

7.2 Suggestions for Future Work

It would be interesting to consider one or more defenders which defend the targets by attacking the vehicles. This would extend the problem investigated in this thesis to the pursuit-evasion framework, where the vehicles need to account for the risk of being hit by the defenders in their decision-making process. As another extension to the problem investigated in this thesis, one can consider a zone that the targets aim to enter and the vehicles are to protect by attacking the targets approaching it. In some applications, the targets can only be visited in certain time intervals due to different constraints such as limited availability of targets or time-sensitivity of visiting targets. Considering a specific time window for each target during which the vehicles are allowed would also be an important extension of the present problem statement. Moreover, in a practical setting, there are some limitations in terms of energy consumption of vehicles, there communication and sensing ranges, memory size, computational capability. Some of such limitations, can be addressed using a distributed decision-making strategy. In addition, sometimes different targets may not have the same level of importance. Also sometimes the targets may become more important when they are in certain regions in the mission space. Prioritizing different targets or regions in the mission space can be formulated by using appropriate weighting functions, which can be time-dependent in the most general case.

It would also be important to investigate the case where the locations of the targets are not known. This type of problem arises, for example, in search and rescue operations. In an adversarial environment, one the other hand, it may not be possible to guarantee the elimination of the targets. One can use a probabilistic framework to formulate this type of scenario, by considering a probability of success during the vehicle-target engagement. Moreover, in an uncertain environment and also in the case where the communications and sensing signals are prone to noise, it would be of practical importance to consider the problems such as false alarms, soft attacks and jamming. In all of the problems discussed above, a receding horizon approach similar to the one proposed in this thesis can be most effective.

Bibliography

[1] R. M. Murray, "Recent research in cooperative control of multivehicle systems," *Journal of Dynamic Systems, Measurement, and Control*, vol. 129, no. 5, pp. 571–583, 2007.

[2] F. Bullo, J. Cortés, and S. Martinez, *Distributed Control of Robotic Networks: A Mathematical Approach to Motion Coordination Algorithms: A Mathematical Approach to Motion Coordination Algorithms.* Princeton University Press, 2009.

[3] R. J. Wood, "The first takeoff of a biologically inspired at-scale robotic insect," *Robotics, IEEE Transactions on*, vol. 24, no. 2, pp. 341–347, 2008.

[4] T. Clutton-Brock, "Cooperation between non-kin in animal societies," *Nature*, vol. 462, no. 7269, pp. 51–57, 2009.

[5] M. A. Nowak, "Five rules for the evolution of cooperation," *science*, vol. 314, no. 5805, pp. 1560–1563, 2006.

[6] L. Mui, M. Mohtashemi, and A. Halberstadt, "A computational model of trust and reputation," in *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on.* IEEE, 2002, pp. 2431–2439.

[7] V. Srinivasan, P. Nuggehalli, C. F. Chiasserini, and R. R. Rao, "Cooperation in wireless ad hoc networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 2. IEEE, 2003, pp. 808–817.

[8] N. T. Bailey *et al.*, *The mathematical theory of infectious diseases and its applications.* Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE., 1975.

[9] C. Castelfranchi and R. Falcone, "Principles of trust for mas: Cognitive anatomy, social importance, and quantification," in *Multi Agent Systems, 1998. Proceedings. International Conference on.* IEEE, 1998, pp. 72–79.

[10] N. R. Jennings, K. Sycara, and M. Wooldridge, "A roadmap of agent research and development," *Autonomous agents and multi-agent systems*, vol. 1, no. 1, pp. 7–38, 1998.

[11] M. Wooldridge, *An introduction to multiagent systems.* John Wiley & Sons, 2009.

[12] J. Ferber, *Multi-agent systems: an introduction to distributed artificial intelligence.* Addison-Wesley Reading, 1999, vol. 1.

[13] C. Nowzari, V. M. Preciado, and G. J. Pappas, "Analysis and control of epidemics: A survey of spreading processes on complex networks," *Control Systems, IEEE*, vol. 36, no. 1, pp. 26–46, 2016.

[14] Y. Wan, S. Roy, and A. Saberi, "Designing spatially heterogeneous strategies for control of virus spread," *Systems Biology, IET*, vol. 2, no. 4, pp. 184–201, 2008.

[15] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. Pappas, "Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks," in *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*. IEEE, 2013, pp. 7486–7491.

[16] E. Ramirez-Llanos and S. Martinez, "A distributed algorithm for virus spread minimization," in *American Control Conference (ACC), 2014*. IEEE, 2014, pp. 184–189.

[17] Y. Hayel, S. Trajanovski, E. Altman, H. Wang, and P. Van Mieghem, "Complete game-theoretic characterization of sis epidemics protection strategies," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*. IEEE, 2014, pp. 1179–1184.

[18] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *Automatic Control, IEEE Transactions on*, vol. 49, no. 9, pp. 1520–1533, 2004.

[19] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *Automatic Control, IEEE Transactions on*, vol. 51, no. 3, pp. 401–420, 2006.

[20] R. Olfati-Saber, A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[21] A. Jadbabaie, J. Lin *et al.*, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *Automatic Control, IEEE Transactions on*, vol. 48, no. 6, pp. 988–1001, 2003.

[22] M. Mesbahi and M. Egerstedt, *Graph theoretic methods in multiagent networks.* Princeton University Press, 2010.

[23] A. Khamis, A. Hussein, and A. Elmogy, "Multi-robot task allocation: A review of the state-of-the-art," in *Cooperative Robots and Sensor Networks 2015.* Springer, 2015, pp. 31–51.

[24] L. E. Parker, "Alliance: An architecture for fault tolerant multirobot cooperation," *Robotics and Automation, IEEE Transactions on*, vol. 14, no. 2, pp. 220–240, 1998.

[25] W. Burgard, M. Moors, C. Stachniss, and F. E. Schneider, "Coordinated multirobot exploration," *Robotics, IEEE Transactions on*, vol. 21, no. 3, pp. 376–386, 2005.

[26] A. K. Das, R. Fierro, V. Kumar, J. P. Ostrowski, J. Spletzer, and C. J. Taylor, "A vision-based formation control framework," *Robotics and Automation, IEEE Transactions on*, vol. 18, no. 5, pp. 813–825, 2002.

[27] T. Balch and R. C. Arkin, "Behavior-based formation control for multirobot teams," *Robotics and Automation, IEEE Transactions on*, vol. 14, no. 6, pp. 926–939, 1998.

[28] M. Garetto, W. Gong, and D. Towsley, "Modeling malware spreading dynamics," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 3. IEEE, 2003, pp. 1869–1879.

[29] R. Johari and D. K. H. Tan, "End-to-end congestion control for the internet: Delays and stability," *Networking, IEEE/ACM Transactions on*, vol. 9, no. 6, pp. 818–832, 2001.

[30] A. Ajorlou, A. Jadbabaie, and A. Kakhbod, "Dynamic pricing in social networks: The word of mouth effect," *Available at SSRN 2495509*, 2015.

[31] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[32] B. Grocholsky, J. Keller, V. Kumar, and G. Pappas, "Cooperative air and ground surveillance," *Robotics & Automation Magazine, IEEE*, vol. 13, no. 3, pp. 16–25, 2006.

[33] V. K. Singh, P. K. Atrey, and M. S. Kankanhalli, "Coopetitive multi-camera surveillance using model predictive control," *Machine Vision and applications*, vol. 19, no. 5-6, pp. 375–393, 2008.

[34] J. S. Jennings, G. Whelan, and W. F. Evans, "Cooperative search and rescue with a team of mobile robots," in *Advanced Robotics, 1997. ICAR'97. Proceedings., 8th International Conference on.* IEEE, 1997, pp. 193–200.

[35] R. R. Murphy, "Human-robot interaction in rescue robotics," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 34, no. 2, pp. 138–153, 2004.

[36] Y. Liu and G. Nejat, "Robotic urban search and rescue: A survey from the control perspective," *Journal of Intelligent & Robotic Systems*, vol. 72, no. 2, pp. 147–165, 2013.

[37] Y. Bao, X. Fu, and X. Gao, "Path planning for reconnaissance uav based on particle swarm optimization," in *Computational Intelligence and Natural Computing Proceedings (CINC), 2010 Second International Conference on*, vol. 2. IEEE, 2010, pp. 28–32.

[38] P. Dasgupta, "A multiagent swarming system for distributed automatic target recognition using unmanned aerial vehicles," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 38, no. 3, pp. 549–563, 2008.

[39] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *Communications magazine, IEEE*, vol. 40, no. 8, pp. 102–114, 2002.

[40] ——, "Wireless sensor networks: a survey," *Computer networks*, vol. 38, no. 4, pp. 393–422, 2002.

[41] C. Desjardins and B. Chaib-draa, "Cooperative adaptive cruise control: A reinforcement learning approach," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 12, no. 4, pp. 1248–1260, 2011.

[42] T. B. Curtin, J. G. Bellingham, J. Catipovic, and D. Webb, "Autonomous oceanographic sampling networks," *Oceanography*, vol. 6, no. 3, pp. 86–94, 1993.

[43] N. E. Leonard, D. A. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni, and R. E. Davis, "Collective motion, sensor networks, and ocean sampling," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 48–74, 2007.

[44] S. Martinez, J. Cortes, and F. Bullo, "Motion coordination with distributed information," *Control Systems, IEEE*, vol. 27, no. 4, pp. 75–88, 2007.

[45] N. Moshtagh, N. D. Michael, A. Jadbabaie, and K. Daniilidis, "Vision-based, distributed control laws for motion coordination of nonholonomic robots," 2009.

[46] W. Ren, R. W. Beard, and E. M. Atkins, "A survey of consensus problems in multi-agent coordination," in *American Control Conference, 2005. Proceedings of the 2005.* IEEE, 2005, pp. 1859–1864.

[47] R. W. Beard, J. Lawton, F. Y. Hadaegh *et al.*, "A coordination architecture for spacecraft formation control," *IEEE Transactions on control systems technology*, vol. 9, no. 6, pp. 777–790, 2001.

[48] M. Mesbahi and F. Y. Hadaegh, "Formation flying control of multiple spacecraft via graphs, matrix inequalities, and switching," *Journal of Guidance, Control, and Dynamics*, vol. 24, no. 2, pp. 369–377, 2001.

[49] C. Tomlin, G. J. Pappas, and S. Sastry, "Conflict resolution for air traffic management: A study in multiagent hybrid systems," *Automatic Control, IEEE Transactions on*, vol. 43, no. 4, pp. 509–521, 1998.

[50] A. R. Girard, A. S. Howell, and J. K. Hedrick, "Border patrol and surveillance missions using multiple unmanned air vehicles," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 1. IEEE, 2004, pp. 620–625.

[51] A. Ajorlou, A. Momeni, and A. G. Aghdam, "A class of bounded distributed control strategies for connectivity preservation in multi-agent systems," *Automatic Control, IEEE Transactions on*, vol. 55, no. 12, pp. 2828–2833, 2010.

[52] S. Rathinam, R. Sengupta, and S. Darbha, "A resource allocation algorithm for multivehicle systems with nonholonomic constraints," *Automation Science and Engineering, IEEE Transactions on*, vol. 4, no. 1, pp. 98–104, 2007.

[53] T. Maddula, A. A. Minai, and M. M. Polycarpou, "Multi-target assignment and path planning for groups of uavs," in *Recent Developments in Cooperative Control and Optimization*. Springer, 2004, pp. 261–272.

[54] R. W. Beard, T. W. McLain, M. Goodrich, E. P. Anderson *et al.*, "Coordinated target assignment and intercept for unmanned air vehicles," *Robotics and Automation, IEEE Transactions on*, vol. 18, no. 6, pp. 911–922, 2002.

[55] T. H. Chung, G. A. Hollinger, and V. Isler, "Search and pursuit-evasion in mobile robotics," *Autonomous robots*, vol. 31, no. 4, pp. 299–316, 2011.

[56] C. Robin and S. Lacroix, "Multi-robot target detection and tracking: taxonomy and survey," *Autonomous Robots*, pp. 1–32, 2015.

[57] P. J. Nahin, *Chases and escapes: the mathematics of pursuit and evasion*. Princeton University Press, 2012.

[58] L. Angelani, "Collective predation and escape strategies," *Physical review letters*, vol. 109, no. 11, p. 118104, 2012.

[59] T. Ishida and R. E. Korf, "Moving target search." in *IJCAI*, vol. 91, 1991, pp. 204–210.

[60] R. Nowakowski and P. Winkler, "Vertex-to-vertex pursuit in a graph," *Discrete Mathematics*, vol. 43, no. 2-3, pp. 235–239, 1983.

[61] J. E. Littlewood, *A mathematicians's miscellany.* Methuen, 1953.

[62] H. T. Croft, "lion and man: A postscript," *Journal of the London Mathematical Society*, vol. 1, no. 1, pp. 385–390, 1964.

[63] D. Shasha, J. T. Wang, and R. Giugno, "Algorithmics and applications of tree and graph searching," in *Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems.* ACM, 2002, pp. 39–52.

[64] I. Rusnak, "The lady, the bandit and the body-guard game," in *The 44th Israel Annual Conference on Aerospace Science*, 2004.

[65] ——, "The lady, the bandits, and the bodyguards–a two team dynamic game," in *Proceedings of the 16th World IFAC Congress*, 2005.

[66] E. Garcia, D. W. Casbeer, K. Pham, and M. Pachter, "Cooperative aircraft defense from an attacking missile," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on.* IEEE, 2014, pp. 2926–2931.

[67] F. V. Fomin and D. M. Thilikos, "An annotated bibliography on guaranteed graph searching," *Theoretical computer science*, vol. 399, no. 3, pp. 236–245, 2008.

[68] A. Kolling and S. Carpin, "Pursuit-evasion on trees by robot teams," *IEEE Transactions on Robotics*, vol. 26, no. 1, pp. 32–47, 2010.

[69] Y. Yavin and M. Pachter, *Pursuit-evasion differential games.* Elsevier, 2014, vol. 14.

[70] R. Vidal, O. Shakernia, H. J. Kim, D. H. Shim, and S. Sastry, "Probabilistic pursuit-evasion games: theory, implementation, and experimental evaluation," *Robotics and Automation, IEEE Transactions on*, vol. 18, no. 5, pp. 662–669, 2002.

[71] S. D. Bopardikar, F. Bullo, and J. P. Hespanha, "Cooperative pursuit with sensing limitations," in *American Control Conference, 2007. ACC'07.* IEEE, 2007, pp. 5394–5399.

[72] D. Li, J. B. Cruz, and C. J. Schumacher, "Stochastic multi-player pursuit-evasion differential games," *International Journal of Robust and Nonlinear Control*, vol. 18, no. 2, p. 218, 2008.

[73] S. D. Bopardikar, F. Bullo, and J. P. Hespanha, "A cooperative homicidal chauffeur game," *Automatica*, vol. 45, no. 7, pp. 1771–1777, 2009.

[74] M. Wei, G. Chen, J. B. Cruz, L. Haynes, K. Pham, and E. Blasch, "Multi-pursuer multi-evader pursuit-evasion games with jamming confrontation," *Journal of Aerospace Computing, Information, and Communication*, vol. 4, no. 3, pp. 693–706, 2007.

[75] D. Li, J. B. Cruz Jr, G. Chen, C. Kwan, and M.-H. Chang, "A hierarchical approach to multi-player pursuit-evasion differential games," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on.* IEEE, 2005, pp. 5674–5679.

[76] G. Laporte, "The traveling salesman problem: An overview of exact and approximate algorithms," *European Journal of Operational Research*, vol. 59, no. 2, pp. 231–247, 1992.

[77] C. H. Papadimitriou, "The euclidean travelling salesman problem is np-complete," *Theoretical Computer Science*, vol. 4, no. 3, pp. 237–244, 1977.

[78] A. Orman and H. P. Williams, "A survey of different integer programming formulations of the travelling salesman problem," *Optimisation, economics and financial analysis. Advances in computational management science*, vol. 9, pp. 93–106, 2006.

[79] H. Kaplan, M. Lewenstein, N. Shafrir, and M. Sviridenko, "Approximation algorithms for asymmetric tsp by decomposing directed regular multigraphs," *Journal of the ACM (JACM)*, vol. 52, no. 4, pp. 602–626, 2005.

[80] E. K. Baker, "Technical notean exact algorithm for the time-constrained traveling salesman problem," *Operations Research*, vol. 31, no. 5, pp. 938–945, 1983.

[81] D. Feillet, P. Dejax, and M. Gendreau, "Traveling salesman problems with profits," *Transportation science*, vol. 39, no. 2, pp. 188–205, 2005.

[82] P. Vansteenwegen, W. Souffriau, and D. Van Oudheusden, "The orienteering problem: A survey," *European Journal of Operational Research*, vol. 209, no. 1, pp. 1–10, 2011.

[83] T. Bektas, "The multiple traveling salesman problem: an overview of formulations and solution procedures," *Omega*, vol. 34, no. 3, pp. 209–219, 2006.

[84] B. L. Golden, S. Raghavan, and E. A. Wasil, *The vehicle routing problem: latest advances and new challenges.* Springer Science & Business Media, 2008, vol. 43.

[85] G. Laporte, "The vehicle routing problem: An overview of exact and approximate algorithms," *European Journal of Operational Research*, vol. 59, no. 3, pp. 345–358, 1992.

[86] H. Min, "The multiple vehicle routing problem with simultaneous delivery and pick-up points," *Transportation Research Part A: General*, vol. 23, no. 5, pp. 377–386, 1989.

[87] B. Kallehauge, J. Larsen, O. B. Madsen, and M. M. Solomon, *Vehicle routing problem with time windows.* Springer, 2005.

[88] P. Toth and D. Vigo, "Models, relaxations and exact approaches for the capacitated vehicle routing problem," *Discrete Applied Mathematics*, vol. 123, no. 1, pp. 487–512, 2002.

[89] J. Brandao and A. Mercer, "The multi-trip vehicle routing problem," *Journal of the Operational research society*, pp. 799–805, 1998.

[90] J. Brandão, "A tabu search algorithm for the open vehicle routing problem," *European Journal of Operational Research*, vol. 157, no. 3, pp. 552–564, 2004.

[91] V. Pillac, M. Gendreau, C. Guéret, and A. L. Medaglia, "A review of dynamic vehicle routing problems," *European Journal of Operational Research*, vol. 225, no. 1, pp. 1–11, 2013.

[92] F. Bullo, E. Frazzoli, M. Pavone, K. Savla, and S. L. Smith, "Dynamic vehicle routing for robotic systems," *Proceedings of the IEEE*, vol. 99, no. 9, pp. 1482–1504, 2011.

[93] D. J. Bertsimas and G. Van Ryzin, "A stochastic and dynamic vehicle routing problem in the euclidean plane," *Operations Research*, vol. 39, no. 4, pp. 601–615, 1991.

[94] ——, "Stochastic and dynamic vehicle routing with general demand and interarrival time distributions," *Advances in Applied Probability*, pp. 947–978, 1993.

[95] ——, "Stochastic and dynamic vehicle routing in the euclidean plane with multiple capacitated vehicles," *Operations Research*, vol. 41, no. 1, pp. 60–76, 1993.

[96] H. Psaraftis, "Dynamic vehicle routing problems. bl golden, aa assad, eds," *Vehicle Routing: Methods and Studies*, vol. 16, pp. 223–248, 1988.

[97] M. Pavone, E. Frazzoli, and F. Bullo, "Adaptive and distributed algorithms for vehicle routing in a stochastic and dynamic environment," *Automatic Control, IEEE Transactions on*, vol. 56, no. 6, pp. 1259–1274, 2011.

[98] M. Hammar and B. J. Nilsson, "Approximation results for kinetic variants of tsp," *Discrete & Computational Geometry*, vol. 27, no. 4, pp. 635–651, 2002.

[99] S. D. Bopardikar, S. L. Smith, F. Bullo, and J. P. Hespanha, "Dynamic vehicle routing for translating demands: Stability analysis and receding-horizon policies," *Automatic Control, IEEE Transactions on*, vol. 55, no. 11, pp. 2554–2569, 2010.

[100] P. Chalasani and R. Motwani, "Approximating capacitated routing and delivery problems," *SIAM Journal on Computing*, vol. 28, p. 1995, 1995.

[101] Y. Asahiro, E. Miyano, and S. Shimoirisa, "Grasp and delivery for moving objects on broken lines," *Theory of Computing Systems*, vol. 42, no. 3, pp. 289–305, 2008.

[102] S. D. Bopardikar, S. L. Smith, and F. Bullo, "On dynamic vehicle routing with time constraints," *Robotics, IEEE Transactions on*, vol. 30, no. 6, pp. 1524–1532, 2014.

[103] P. Agharkar and F. Bullo, "Vehicle routing algorithms to intercept escaping targets," in *American Control Conference (ACC), 2014*. IEEE, 2014, pp. 952–957.

[104] P. Agharkar, S. D. Bopardikar, and F. Bullo, "Vehicle routing algorithms for radially escaping targets," *SIAM Journal on Control and Optimization*, vol. 53, no. 5, pp. 2934–2954, 2015.

[105] M. Zhong and C. G. Cassandras, "Distributed coverage control and data collection with mobile sensor networks," *Automatic Control, IEEE Transactions on*, vol. 56, no. 10, pp. 2445–2455, 2011.

[106] G. A. Hollinger, S. Choudhary, P. Qarabaqi, C. Murphy, U. Mitra, G. S. Sukhatme, M. Stojanovic, H. Singh, and F. Hover, "Underwater data collection using robotic sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 5, pp. 899–911, 2012.

[107] C. G. Cassandras and W. Li, "A receding horizon approach for solving some cooperative control problems," in *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, vol. 4. IEEE, 2002, pp. 3760–3765.

[108] W. Li and C. G. Cassandras, "A cooperative receding horizon controller for multivehicle uncertain environments," *Automatic Control, IEEE Transactions on*, vol. 51, no. 2, pp. 242–257, 2006.

[109] ——, "Centralized and distributed cooperative receding horizon control of autonomous vehicle missions," *Mathematical and computer modelling*, vol. 43, no. 9, pp. 1208–1228, 2006.

[110] E. Balas, "The prize collecting traveling salesman problem," *Networks*, vol. 19, no. 6, pp. 621–636, 1989.

[111] A. Ekici and A. Retharekar, "Multiple agents maximum collection problem with time dependent rewards," *Computers & Industrial Engineering*, vol. 64, no. 4, pp. 1009–1018, 2013.

[112] H. Tang, E. Miller-Hooks, and R. Tomastik, "Scheduling technicians for planned maintenance of geographically distributed equipment," *Transportation Research Part E: Logistics and Transportation Review*, vol. 43, no. 5, pp. 591–609, 2007.

[113] Y. Khazaeni and C. G. Cassandras, "A new event-driven cooperative receding horizon controller for multi-agent systems in uncertain environments," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*. IEEE, 2014, pp. 2770–2775.

[114] H. M. Pari, M. Khosravi, and M. Haeri, "Model predictive based dynamic path planning for single target tracking and formation," in *Control, Instrumentation, and Automation (ICCIA), 2013 3rd International Conference on*. IEEE, 2013, pp. 340–344.

[115] E. Bakolas and P. Tsiotras, "Optimal pursuit of moving targets using dynamic voronoi diagrams," in *Decision and Control (CDC), 2010 49th IEEE Conference on*. IEEE, 2010, pp. 7431–7436.

[116] Y. Lan, "Multiple mobile robot cooperative target intercept with local coordination," in *Control and Decision Conference (CCDC), 2012 24th Chinese*. IEEE, 2012, pp. 145–151.

[117] J. Bellingham, M. Tillerson, A. Richards, and J. P. How, "Multi-task allocation and path planning for cooperating uavs," in *Cooperative control: models, applications and algorithms*. Springer, 2003, pp. 23–41.

[118] M. G. Earl and R. DAndrea, "Multi-vehicle cooperative control using mixed integer linear programming," *Cooperative Control of Distributed Multi-Agent Systems*, pp. 231–259, 2007.

[119] M. G. Earl and R. D Andrea, "Modeling and control of a multi-agent system using mixed integer linear programming," in *IEEE Conference on Decision and Control*, vol. 1. IEEE; 1998, 2001, pp. 107–111.

[120] M. Khosravi and A. G. Aghdam, "Stability analysis of dynamic decision-making for vehicle heading control," in *American Control Conference (ACC)*. IEEE, 2015, pp. 3076–3081.

[121] M. Khosravi, H. Khodadadi, H. Rivaz, and A. G. Aghdam, "Cooperative control for multi-target interception with sensing and communication limitations: A game theoretic approach," in *Decision and Control (CDC), accepted*. IEEE, 2015.

[122] M. Khosravi and A. G. Aghdam, "Cooperative receding horizon control for multi-target interception in uncertain environments," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*. IEEE, 2014, pp. 4497–4502.

[123] M. Khosravi, H. Khodadadi, H. Rivaz, and A. G. Aghdam, "Cooperative Receding Horizon Control of Double Integrator Vehicles for Multi-Target Interception ," in *American Control Conference (ACC)*. IEEE, 2016.

[124] M. Khosravi, H. Khodadadi, A. G. Aghdam, and H. Rivaz, "Maximum reward collection problem: a cooperative receding horizon approach for dynamic clustering," in *Proceedings of the 2015 Conference on Research in Adaptive and Convergent Systems*. ACM, 2015, pp. 38–43.

[125] M. M. Asadi, M. Khosravi, A. G. Aghdam, and S. Blouin, "Joint power optimization and connectivity control problem over underwater random sensor networks." IEEE, 2015, pp. 2709–2714.

[126] ——, "Generalized Algebraic Connectivity for Asymmetric Networks ," in *American Control Conference (ACC)*. IEEE, 2016.

[127] M. Kline, *Calculus: an intuitive and physical approach*. Courier Corporation, 1998.

[128] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[129] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear programming: theory and algorithms*. John Wiley & Sons, 2013.

[130] W. Rudin, *Principles of mathematical analysis*. McGraw-Hill New York, 1964, vol. 3.

[131] G. Grimmett and D. Stirzaker, *Probability and random processes*. Oxford university press, 2001.

[132] G. Arslan, J. R. Marden, and J. S. Shamma, "Autonomous vehicle-target assignment: A game-theoretical formulation," *Journal of Dynamic Systems, Measurement, and Control*, vol. 129, no. 5, pp. 584–596, 2007.

[133] G. Arslan and J. S. Shamma, "Distributed vehicle-target assignment using learning in games," in *Decision and Control, 2006 45th IEEE Conference on*. IEEE, 2006, pp. 2837–2842.

[134] A. Arsie, K. Savla, and E. Frazzoli, "Efficient routing algorithms for multiple vehicles with no explicit communications," *Automatic Control, IEEE Transactions on*, vol. 54, no. 10, pp. 2302–2317, 2009.