Abstract—In this paper, the cooperative multi-target interception problem in uncertain environment with double-integrator vehicles is investigated. By defining a time discounting reward function for each target which can be collected only if it is visited by a vehicle, this problem can be formulated as an optimization problem which maximizes the expected reward collectible from the set of available targets in the mission space. 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. Assuming double-integrator dynamics for the vehicles further complicates the problem due to the impact of inertia which penalizes changes in target assignments. Despite these issues, the performance of the proposed algorithm is assessed and analytically confirmed. Moreover, the effectiveness and advantages of the proposed algorithm is demonstrated via numerical simulations.

I. INTRODUCTION

In recent years, cooperative control of multi-agent systems [1], [2] has attracted substantial research interest in the variety of applications such as surveillance [3], search and rescue missions [4] formation [5], consensus [6], and target tracking and vehicles target assignment [7], to name only a few. The idea behind cooperative control of multi-agent systems is to fulfill a global objective with a set of simple agents with limited ability and the proper use of information exchange among them.

Multi-target interception problem with multiple vehicles is one of the more recently touched topics in cooperative control systems where multiple vehicles are cooperatively planning to visit targets that appear in the mission space at random time instants [7]. In [8], cooperative multi-target interception problem is solved, with no a priori knowledge about the arrival times of the target points, utilizing a cooperative receding horizon (CRH) control scheme with the sum assignment. In [9], the authors further improved the controller developed in [8] and overcame some of its limitations such as poor performance and instability in target trajectories. In [10], the vehicles are assigned to the moving objects with known kinematics based on dynamic Voronoi partitioning. In [11], a receding horizon control approach is proposed for multi-target interception by a single vehicle in an uncertain environment. In [12], a cooperative receding horizon controller is designed for intercepting a set of target moving in the mission space on a priori unknown trajectories. This problem is extended in [13] to a decentralized setting in which the agents have limited sensing and communication capabilities and is solved using a game-theoretic approach. Dynamic vehicle routing (DVR) is a similar problem to multi-target interception and aims at minimizing the time of target tracking and servicing demands is investigated in [14]. In [15], a DVR problem is studied for targets which appear based on a Poisson process, and then move outward of the mission space (unit disk) with a constant speed.

However, in all of the aforementioned papers, there are some restrictive assumptions: the targets are assumed to be stationary points in [8], [9], [14], only a single vehicle is considered for mission accomplishment in [11], [15], and the arrival time of the targets are assumed to be known in [10]. More importantly taking into account the fact that a broad class of vehicles requires a double-integrator dynamic model, i.e., mass-force model, single integrator vehicles [8]–[15] cannot be a good representative of the real applications. However, assuming double-integrator dynamics for the vehicles further complicates the problem due to the impact of inertia which penalize changes in target assignments.

In this paper, the cooperative multi-target interception problem in uncertain environment with double-integrator vehicles is investigated. Similar to [8], [9], [11]–[13], 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.

The paper is structured as follows: In Section II, the problem formulation, consisting of definitions and assumptions, is presented. Section III discusses the reformulation of target interception problem to the maximum reward collection problem, structure of cooperation strategy and cooperative receding horizon controller. Section IV deals with stationary analysis of cooperative receding horizon controller. In section V an illustrative simulation is presented and the paper is concluded.
in section VI.

A. 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 like \( B \), the indicator function of \( B \) is denoted by \( 1_B \) which is a function from \( A \) to \{0, 1\} where it is non-zero only when its argument belong to the set \( B \). For any index set \( I \), the \( A^I \) represents the set of points like \((a_i)_{i \in I}\) where each of its entries belongs to \( A \). In the case that \( I \) is the set \( \mathbb{N}_n \), the \( A^\mathbb{N}_n \) is simply shown by \( A^n \).

Let \( R_j \) and \( v_j \) be the state and velocity vector of vehicle \( j \), \( d_j \) and \( q_j \) be the dynamics described as following the control input for its magnitude, for any \( j \in \mathcal{I}_V \) and any given time \( t \in \mathbb{R}_{\geq 0} \).

Let \( \mathcal{I}_T = N_{\mathcal{I}_T} \) be a finite set of natural numbers representing indices of a non-zero finite number of targets sequentially arriving the mission space. Assume that the mission starts at time \( t = 0 \), and let \( n_0 \in (0) \cup N_{\mathcal{I}_T} \) be the number of targets in the mission space at the start of the mission. Let \( T_1 \) be the arrival time of the first target, and \( (T_i)_{i=2}^n \) be a finite sequence of non-negative real scalars representing targets interarrival times, the time between consecutive targets arrival. Note that if \( n_0 > 0 \), then for any \( 1 \leq i \leq n_0 \), one has \( T_i = 0 \). Based on the interarrival times, for any \( i \in N_{\mathcal{I}_T} \), one can define the arrival time of \( i^{th} \) target as \( \tau_i = \sum_{j=1}^{i} T_j \), and also the set of indices of targets arrived up to time moment \( t \), denoted by \( \hat{\mathcal{I}_T}(t) \), as

\[
\hat{\mathcal{I}_T}(t) := \{ i \in \mathcal{I}_T ; \tau_i \leq t \}.
\]  

It is worth noting that \( \{\tau_i\}_{i \in \mathcal{I}_T} \) is an increasing finite sequence. Besides the arrival time of targets, one can set a sequence of vectors like \( \{\bar{r}_i\}_{i \in \mathcal{I}_T} \), belonging to \( \mathcal{M} \), as the initial position of targets in the mission space as they arrive.

**Assumption 1:** The arrival times and initial positions of targets are a priori unknown. More precisely, at any time \( t \in \mathbb{R}_{\geq 0} \) such that \( t < \tau_i \), none of the vehicles has the information of \( \tau_i \) and \( r_i \).

Having these all, one can say that for any \( i \in \mathcal{I}_T \), the \( i^{th} \) target arrives in the mission space at a priori unknown time \( \tau_i \) and in a priori unknown point \( r_i \), and also moves on an a priori unknown trajectory, denoted by \( r_i(t) \).

**Definition 1:** Let \( i \in \mathcal{I}_T \) and \( j \in \mathcal{I}_V \). Given a positive scalar \( d_{ij} \), it is said that the \( j^{th} \) vehicle visits the \( i^{th} \) target at time \( t \), if \( \|p_j(t) - r_i(t)\| \leq d_{ij} \).

**Remark 2:** The reason behind the introduction of scalar \( d_{ij} \) in Definition 1 is to capture practical considerations regarding the physical size of the targets and vehicles. Indeed, if each target and each vehicle is expressed as a point of mass, the scalar \( d_{ij} \) is used to account for the size of the \( i^{th} \) target and the size of the \( j^{th} \) vehicle. For example, for any \( i \in \mathcal{I}_T \) and any \( j \in \mathcal{I}_V \), if target \( i \) and vehicle \( j \) have approximately spherical shapes, respectively with radius \( r_i \) and \( s_j \), then one can set \( d_{ij} := r_i + s_j \).

Along with Definition 1, for any \( i \in \mathcal{I}_T \), one can define \( \hat{\tau}_i \in \mathbb{R}_{\geq 0} = [0, \infty) \) as the first time that target \( i \) is visited by one of the vehicles, i.e.

\[
\hat{\tau}_i = \inf \{ t \in \mathbb{R}_{\geq 0} ; \min_{j \in \mathcal{I}_V} (\|p_j(t) - r_i(t)\| - d_{ij}) \leq 0 \}.
\]  

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 \( \hat{\mathcal{I}_T}(t) \), is defined as following

\[
\hat{\mathcal{I}_T}(t) := \{ i \in \mathcal{I}_T ; \hat{\tau}_i \leq t \}.
\]  

Also one can define the set of indices of targets arrived in the mission space but not visited up to time \( t \), denoted by \( \tilde{\mathcal{I}_T}(t) \), as following

\[
\tilde{\mathcal{I}_T}(t) = \hat{\mathcal{I}_T}(t) \setminus \hat{\mathcal{I}_T}(t) = \{ i \in \mathcal{I}_T ; \hat{\tau}_i \leq t < \hat{\tau}_i \}.
\]
For any \( i \in \mathcal{I}_T \), the trajectory of target \( i \) is a \( C^2 \) curve in the mission space \( \mathcal{M} \), defined as a function \( r_i : [\hat{\tau}_i, \tilde{\tau}_i] \rightarrow \mathcal{M} \) satisfying the following two geometric conditions where one describes global behavior of trajectories of targets and the other one describes local behavior of trajectories of targets.

**Assumption 2:** (Global Geometric Condition) For any \( i \in \mathcal{I}_T \) and any \( \tau \in [\hat{\tau}_i, \tilde{\tau}_i] \), one has that \( r_i(\tau) \in \mathcal{M} \). By global geometric condition, which is guaranteed that targets which have arrived will remain inside the mission space until the end of the mission. Note that the property stated in Assumption 2 depends both on trajectories of targets and also on the geometry of the mission space. For example, in the case that \( \mathcal{M} \) is the whole \( d \)-dimensional Euclidean space, one can easily see that the global geometric condition is satisfied automatically.

**Assumption 3:** (Local Geometric Condition) For any \( i \in \mathcal{I}_T \), one has that \( r_i : [\hat{\tau}_i, \tilde{\tau}_i] \rightarrow \mathcal{M} \) is a \( C^2 \) function, i.e. \( r_i \) is two times continuously differentiable. Also, there exist non-negative scalars \( \hat{a}_i, \hat{h}_i \) such that for any \( i \in \mathcal{I}_T \), \( \tau \in [\hat{\tau}_i, \tilde{\tau}_i] \) and \( t \in [\tau, \tilde{\tau}_i] \), one has
\[
\| \frac{d^2}{dt^2} r_i(\tau) \| \leq \hat{a}_i, \quad \sup_{t \in [\tau, \tilde{\tau}_i]} \| j_i(\tau, \tau) \| \leq \hat{h}_i, \tag{7}
\]
where \( j_i(t, \tau) \) is a \( C^2 \) function satisfying the following equality
\[
r_i(t) = r_i(\tau) + (t - \tau) \frac{d}{dt} r_i(\tau) + \frac{1}{2!} (t - \tau)^2 \frac{d^2}{dt^2} r_i(\tau)
+ \frac{1}{3!} (t - \tau)^3 j_i(t, \tau). \tag{8}
\]

**Remark 3:** For any \( i \in \mathcal{I}_T \), regarding the function \( j_i \), one can say that \( j_i : \mathcal{R}_{\hat{\tau}_i, \tilde{\tau}_i} \rightarrow \mathbb{R}^d \) where
\[
\mathcal{R}_{\hat{\tau}_i, \tilde{\tau}_i} = \{ (t, \tau) ; \ \hat{\tau}_i \leq \tau < t \leq \tilde{\tau}_i \}, \tag{9}
\]
and it is defined as following
\[
j_i(t, \tau) = \frac{r_i(t) - r_i(\tau) + (t - \tau) \frac{d}{dt} r_i(\tau) + \frac{1}{2!} (t - \tau)^2 \frac{d^2}{dt^2} r_i(\tau)}{\frac{1}{3!} (t - \tau)^3}, \tag{10}
\]
i.e., \( j_i \) can be defined just by using the assumption that \( r_i \) is continuously twice differentiable.

**Assumption 4:** The position, velocity and acceleration vectors of any current target (targets which have been arrived but not visited yet) are available at the beginning of each time horizon (i.e., at time instant \( \tau \) in (8)).

For any \( \tau \in \mathbb{R}_{\geq 0} \) and any \( i \in \mathcal{I}_T(\tau) \), define \( y_i(\tau) \) as the vector of available information of target \( i \) at instant time \( \tau \), i.e.
\[
y_i(\tau) = (r_i(\tau), \frac{d}{dt} r_i(\tau), \frac{d^2}{dt^2} r_i(\tau)). \tag{11}
\]
This information vector belongs to information space of target \( i \) which is denoted by \( \mathcal{Y}_i \) and defined as \( \mathcal{M} \times \mathbb{R}^d \times \mathcal{B}(\mathbb{O}_d, \tilde{a}_i) \).

Accordingly, one can define the information vector of targets at time \( \tau \) as following
\[
y(\tau) = (r(\tau), \frac{d}{dt} r(\tau), \frac{d^2}{dt^2} r(\tau))_{i \in \mathcal{I}_T(\tau)}, \tag{12}
\]
and also, the information space of targets as
\[
\mathcal{Y}_\tau = \times_{i \in \mathcal{I}_T(\tau)} (\mathcal{M} \times \mathbb{R}^d \times \mathcal{B}(\mathbb{O}_d, \tilde{a}_i)). \tag{13}
\]
Considering Assumption 4, for any \( \tau \in \mathbb{R}_{\geq 0} \) and for any \( i \in \mathcal{I}_T(\tau) \), positions 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 moment \( \tau \). Denote this estimation by \( \hat{r}_i(\cdot) \) and set to
\[
\hat{r}_i(t) = r_i(\tau) + (t - \tau) \frac{d}{dt} r_i(\tau) + \frac{1}{2!} (t - \tau)^2 \frac{d^2}{dt^2} r_i(\tau), \tag{14}
\]
where \( t \in [\tau, \tilde{\tau}_i] \).

**Remark 4:** Regarding the precision of this estimation, one can easily see from (14) that
\[
\| \hat{r}_i(t) - r_i(t) \| \leq \frac{1}{3!} (t - \tau)^3 \sup_{t \in [\tau, \tilde{\tau}_i]} \| j_i(t, \tau) \| = \frac{1}{3!} (t - \tau)^3 \hat{y}, \tag{15}
\]
for any \( t \in [\tau, \tilde{\tau}_i] \). Thus, to have a desired precision for estimation (14), resulted from the information available at time instant \( \tau \), one needs to choose \( t \) close enough to \( \tau \).

With respect to each target, a task is defined which is completed if the target is visited by one of the vehicles. By abuse of notation, one can denote by \( \mathcal{I}_T, \hat{T}_T(t), \tilde{T}_T(t) \) and \( \mathcal{T}_T(t) \), as the total set of tasks, set of tasks started by time \( t \), set of tasks accomplished by time \( t \), and the set of tasks in process at time \( t \), respectively. Subsequently, the mission is defined as the operation of accomplishing all of the tasks in a finite time. Here, it is desired to obtain a near-optimal cooperative algorithm for mission accomplishment, considering possible uncertainties and the limitation on information in the presented paradigm.

### III. Cooperative Receding Horizon Scheme

In order to urge the vehicles to accomplish the tasks and accordingly visit the targets, let a time-decreasing reward be assigned to each of the targets which can be collected only if the vehicle completes the corresponding task and visits the target. Vehicles are interested in maximizing the total collected reward, which entails cooperation to minimize the visit time. Toward this goal, the vehicles should decide upon their next immediate targets, in a cooperative manner at each moment of decision-making, and subsequently, plan their own paths. Based on the existence of possible uncertainties in the environment and changes in the given information, the cooperative decision-making and path planning should be done via an iterative procedure. At the beginning of each iteration, the vehicles calculate their control inputs, based on their interests regarding the tasks and their corresponding rewards, such that their estimation of the total collected reward become maximized.

#### A. Structure of Reward Functions

For any \( i \in \mathcal{I}_T \), let \( R_0 \) be the initial reward considered for the task corresponding to the \( i^\text{th} \) target at its arrival moment. In order to take into account the reward loss over time, consider
a continuous decreasing function \( \rho_i: [\tau_i, \tilde{\tau}_i] \rightarrow [0, 1] \), called discount function, and form the reward function as \( \rho_i \cdot r_i \). Assuming that \( \rho_i(\tilde{\tau}_i) = 1 \), 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 target \( i \), one can consider the discount function as follows

\[
\rho_i(t) = e^{-\gamma_i(t-t_i)}, \quad \forall \, i \in \mathcal{I}_T, \tag{16}
\]

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

\[
\rho_i(t) = \max \{1 - \frac{t - \tilde{\tau}_i}{t_f^i - \tilde{\tau}_i}, 0\}, \quad \forall \, i \in \mathcal{I}_T. \tag{17}
\]

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

\[
\rho_i(t) = \begin{cases} 
\rho_{i,0}(t), & \text{if } t \in [\tau_i, D_{i1}), \\
\rho_{i,1}(t), & \text{if } t \in [D_{i1}, D_{i2}), \\
\vdots & \vdots \\
\rho_{i,d}(t), & \text{if } t \in [D_{id}, \tilde{\tau}_i], 
\end{cases} \tag{18}
\]

where \( \{D_{id}\}_{d=1}^D \) are the soft deadlines and \( \{\rho_{i,d}(\cdot)\}_{d=1}^D \) are the discount sub-functions since \( \rho_i(t) \) is continuous function, it is required that for any \( d \in \mathbb{N}_D \), one has \( \lim_{t \rightarrow D_{id}} \rho_{i,d-1}(D_{id}) = \rho_{i,d}(D_{id}) \). 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.

**B. The Minimum Reaching Time and The Maximum Reward Estimation**

Let \( (t_k)_{k=1}^{k_{\max}} \in [0, \tilde{\tau}] \) denote the time instants at which 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 \( t_0 = 0 \) 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 the process as \( I_k = [t_k, t_{k+1}) \).

For any \( j \in \mathcal{I}_V \), from (2), one can represent the dynamic of vehicle \( j \) in matrix form as follows

\[
\frac{d}{dt} x_j = A_d x_j + B_d u_j, \tag{19}
\]

where \( A_d \) and \( B_d \) are \( 2d \times 2d \) matrices defined as

\[
A_d = \begin{pmatrix} 0_d & I_d \\ 0_d & 0_d \end{pmatrix}, \quad B_d = \begin{pmatrix} 0_d \\ I_d \end{pmatrix}. \tag{20}
\]

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

\[
\frac{d}{dt} x = (I_m \otimes A_d) x + (I_m \otimes B_d) u. \tag{21}
\]

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

\[
x_j(t_{k+1}) = e^{A_d(t_{k+1} - t_k)} x_j(t_k) + \int_{t_k}^{t_{k+1}} e^{A_d(t_{k+1} - s)} B_d u_j^k(s) \, ds. \tag{22}
\]

Now, let \( i \in \mathcal{I}_T(t_k) \) be the index of an arbitrary existing target. Let \( \tau_{ij} \) be the time estimate when vehicle \( j \) can reach target \( i \), based on the information given at time instant \( t_k \) and the control input \( u_j^k(\cdot) \) applied by vehicle \( j \) for time interval \( [t_k, t_{k+1}) \). Denote \( \bar{\tau}_i^k, \tilde{\tau}_i^k \) and \( a_i^k \) as \( \tau_{ij}(t_k) = \frac{1}{t_f^k} \bar{\tau}_i^k(t_k)(t_k) \) and \( \frac{1}{t_f^k} \tilde{\tau}_i^k(t_k) \), respectively. From these, one can model the trajectory of target \( i \), for \( t \geq t_k \) as

\[
\dot{i}_i(t) = i_i^k + (t - t_k)v_i^k + \frac{1}{2f}(t - t_k^2a_i^k). \tag{23}
\]

Consequently, one can obtain

\[
\begin{cases}
\dot{i}_i^{k+1} = i_i^{k+1} + (t_{k+1} - t_k)v_i^{k+1} + \frac{1}{2f}(t_{k+1} - t_k^2a_i^{k+1}), \\
\dot{v}_i^{k+1} = v_i^{k+1} + (t_{k+1} - t_k)a_i^{k+1}, \\
\dot{a}_i^{k+1} = a_i^{k+1},
\end{cases} \tag{24}
\]

where \( \bar{\tau}_i^{k+1}, \tilde{\tau}_i^{k+1} \) and \( a_i^{k+1} \) 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 1:** Consider vehicle \( j \in \mathcal{I}_V \), the time instant \( t_k \), the target \( i \in \mathcal{I}_T(t_k) \) and the trajectory model given in (23) for target \( i \). Let the control input \( u_j^k(\cdot) \) be applied for time interval \( I_k \). Also, consider the following equation

\[
\begin{aligned}
\frac{1}{2} u_{ij}^2 \bar{\tau}_{ij}^2 & = \frac{1}{2} \dot{a}_i^{k+1} \bar{\tau}_{ij}^2 + (\dot{v}_i^{k+1} - q_j(t_{k+1})) \bar{\tau}_{ij} \\
\|u_{ij}\| & = u_{\text{max}},
\end{aligned} \tag{25}
\]

where \( \bar{\tau}_i^{k+1}, \tilde{\tau}_i^{k+1} \) and \( a_i^{k+1} \) are the predictions given in (24) and also, \( p_j(t_{k+1}) \) and \( q_j(t_{k+1}) \) are the position and velocity of vehicle \( j \), respectively, given the control input \( u_j^k(\cdot) \) is applied for time interval \( I_k \). Then, equation (25) 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

\[
\frac{1}{2} (u_{\text{max}} + \dot{a}) z^2 = ||(q_j(t_k) - \dot{v}_i^k)z + p_j(t_k) - i_i^k||, \tag{26}
\]

then \( \bar{\tau}_{ij} = t_{k+1} + \tilde{\tau}_{ij} \) where \( \tilde{\tau}_{ij} \) is the smallest non-negative solution of equation (25).

**Proof:** Define continuous function \( f: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^d \) as

\[
f(s) = \dot{a}_i^{k+1} + \frac{2}{s}(\dot{v}_i^{k+1} - q_j(t_{k+1})) + \frac{2}{s^2}(\bar{\tau}_i^{k+1} - p_j(t_{k+1})) \tag{27}
\]
and also the set $\mathcal{Z}$ as \( \{ s \in \mathbb{R}_{>0}; \|f(s)\| = u_{\text{max}} \} \). From the continuity of function \( f \) and norm function, on has that $\mathcal{Z}$ is a closed set. Note that \( \lim_{t \to -\infty} \mathcal{Z} = \mathbb{R}_0^{k+1} \). Since, \( \|\mathcal{Z}\| \leq u_{\text{max}} \), there exists \( \bar{s} \in \mathcal{Z} \). Moreover, it is concluded that $\mathcal{Z}$ is bounded and subsequently, a compact set. Also, if \( \bar{s}_{i+1} - p_i(t_{i+1}) \neq 0 \), one has that \( \lim_{t \to -\infty} \|f(s)\| = 0 \). Therefore, there exists \( \bar{s} \in \mathcal{Z} \). This shows that the compact set $\mathcal{Z}$ is non-empty. Set \( \tau \) as the minimum of $\mathcal{Z}$ and \( u \) as $f(\tau)$. Therefore, \( (\tau, f(\tau)) \) is the solution of (25) with the desired properties.

Now, for any $j \in I_V$ and any $i \in I_T(t_k)$, define function $\Psi_{ij}^{k+1}: \mathbb{R}_{>0} \times \mathbb{R}^{2d} \to \mathbb{R}^d$ as

$$
\Psi_{ij}^{k+1}(t, x) = [I_d, 0_d]x - \frac{1}{2} \hat{\nu}_i^{k+1} t^2 - \hat{v}_i^{k+1} t - \hat{r}_i^{k+1}.
$$

Also, consider the minimum time optimal control problem which is introduced in the following equations:

$$
\begin{align*}
\min_{u} \quad & \int_{0}^{t} f(s) \, ds \\
\text{s.t.} \quad & \frac{d}{dt} w = A_d w + B_d u, \\
& w(0) = x_j(t_{k+1}), \\
& \Psi_{ij}^{k+1}(\tau, w(\tau)) = 0, \\
& u \in \mathcal{U}(u_{\text{max}}).
\end{align*}
$$

Define the Hamiltonian function [16] as

$$
H : \mathbb{R}_{>0} \times \mathbb{R}^{2d} \times \mathbb{R}^{2d} \times \mathbb{R}^d \to \mathbb{R}
$$

with

$$
H(\lambda_0, \lambda, w, u) = \lambda_0 + \lambda^T A_d w + \lambda^T B_d u.
$$

From Pontryagin maximum principle [16], it is concluded that if \( u \) be the optimal solution of (76) and \( w \) be the resulting state trajectory, then there exist a constant $\lambda_0 \in \mathbb{R}_{>0}$ and a costate trajectory $\lambda : [0, \tau] \to \mathbb{R}^{2d}$ which are satisfying the following conditions:

i. For any $t \in [0, \tau]$, one has \( (\lambda_0, \lambda(t)) \neq 0 \).

ii. The costate is a solution to the following differential equation

$$
\frac{d}{dt} \lambda = -A_d^T \lambda.
$$

iii. The following equation holds

$$
H(\lambda_0, \lambda, w, u) = \min_{\nu \in \mathcal{U}} H(\lambda_0, \lambda, w, \nu).
$$

for any $t \in [0, \tau]$. Also, $H(\lambda_0, \lambda, w, u)$ is constant.

iv. There exists vector $\nu \in \mathbb{R}^d$ such that

$$
H = -(D_\nu \Psi_{ij}^{k+1})^T \nu,
$$

and

$$
\lambda(\tau) = (D_\nu \Psi_{ij}^{k+1})^T \nu.
$$

Let $w = [p^T, q^T]^T$ and $\lambda = [\lambda_p^T, \lambda_q^T]^T$. Then, from equation (31) one can see that

$$
\begin{align*}
\frac{d}{dt} \lambda_p &= 0_d, \\
\frac{d}{dt} \lambda_q &= -\lambda_p,
\end{align*}
$$

and therefore, \( \lambda_p \) is constant and \( \lambda_q(t) = \lambda_q(0) - \lambda_p t \), for any $t \in [0, \tau]$. Since $D_q \Psi_{ij}^{k+1}(t, x) = [I_d, 0_d]$, the equation (34) says that there exists vector \( \nu \in \mathbb{R}^d \) such that

$$
[\begin{bmatrix} p(\tau) \\ q(\tau) \end{bmatrix}] = [I_d, 0_d]^T \nu,
$$

and subsequently, $\nu = \lambda_p$ and $\lambda_q(\tau) = 0$. This shows that $\lambda_q(t) = (\tau - t) \lambda_p$. Considering introduced variables $p, q, \lambda_p$ and $\lambda_q$, equation (30) changes to

$$
H(\lambda_0, \lambda_p, \lambda_q, p, q, u) = \lambda_0 + \lambda_q^T q + \lambda_p^T u.
$$

If $\lambda_p = 0_d$, from equation (37) it is concluded that $H = \lambda_0$, and subsequently, from (34) and $\nu = \lambda_p$, one has $\lambda_0 = 0$. Therefore, for any $t \in [0, \tau]$, one has $(\lambda_0, \lambda(t)) = 0$ which contradict with initial assumptions on $\lambda_0$ and $\lambda$. This shows that $\lambda_q \neq 0$. Considering the equation (32) and the fact that $\lambda_q(t) = (\tau - t) \lambda_p$, one can see that the optimal control $u$ is a constant function such that

$$
u(t) = -\frac{\lambda_p}{\|\lambda_p\| u_{\text{max}}}, \quad \forall t \in [0, \tau],
$$

which is here simply denoted by $u$. Given the control input $u$ the state trajectory of system is uniquely defined, specifically $p(\tau)$ and $q(\tau)$ which are as following

$$
p(\tau) = \frac{1}{2} u \tau^2 + q_j(t_{k+1}) + p_j(t_{k+1}),
$$

$$
q(\tau) = u \tau + q_j(t_{k+1}).
$$

Since \( \Psi_{ij}^{k+1}(\tau, w(\tau)) = 0 \), it is required that $u$ satisfies

$$
\frac{1}{2} u \tau^2 + q_j(t_{k+1}) + p_j(t_{k+1}) - \frac{1}{2} \hat{\nu}_i^{k+1} \tau^2 - \hat{v}_i^{k+1} \tau - \hat{r}_i^{k+1} = 0_d,
$$

or equivalently,

$$
\frac{1}{2} u \tau^2 = \frac{1}{2} \hat{\nu}_i^{k+1} \tau^2 + (\hat{v}_i^{k+1} - q_j(t_{k+1})) \tau + \hat{r}_i^{k+1} - p_j(t_{k+1}).
$$

Therefore, $u$ and $\tau$ satisfy the equation (25). From other hand, if $u \in \mathbb{R}^d$ and $\tau \in \mathbb{R}_{>0}$ be the solution of (25) where $\tau$ has the smallest possible value, then one can derive a solution for the optimal control problem satisfying Pontryagin maximum principle. Thus, the equation (25) uniquely determines the solution of the given optimal control problem and determines the minimum reaching time.

Note that the condition introduced in equation (26) guarantees that vehicle $j$ does not visit target $i$ before $t_{k+1}$ (see proof of Lemma 1). Thus, the estimated visit time after $t_k$ will be $t_{k+1} + \tau$. This shows that $t_{ij}^k = t_{k+1} + \tau$ and concludes the proof.

**Corollary 1:** Let the conditions in Theorem 1 hold. Then the maximum reward which vehicle $j$ can collect from target $i$, assuming that the control input $u_i^k(\cdot)$ is applied for the time interval $I_k$, can be estimated as $R \rho_j(t_i^k)$ where $t_i^k$ is the estimation of the reaching time introduced in Theorem 1.

**Remark 5:** Note that at each time instant $t_k$, the minimum reaching time of each target for each of the vehicles, after applying the control inputs $u_i^k$, is an implicit function of
the given information vectors $x^k$ and $y^k$. Subsequently, they are function of the control input $u^k$ and time instant $t_k$, i.e. $\tau_{ij}^k = \pi_{ij}^k(u^k, t_k)$. Similarly, the maximum final rewards are also implicit functions of the given information vectors and accordingly, functions of the control input $u^k$ and time instant $t_k$.

Theorem 1 and Corollary 1 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.

C. 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 $i \in \mathcal{I}_T(t_k)$ and $j \in \mathcal{I}_V$, an assignment of vehicle $j$ to task $i$ is characterized as a real scalar in $[0, 1]$, denoted by $\pi_{ij}^k$, which reflects the amount of interest of vehicle $j$ in being assigned to target $i$ during the time interval $I_k$. Also, denote by $\mathbf{\Pi}^k$ as the assignments matrix which is defined as $(\pi_{ij}^k)_{i \in \mathcal{I}_T(t_k), j \in \mathcal{I}_V}$. It is expected that the value of assignment $\pi_{ij}^k$ 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 $i \in \mathcal{I}_T(t_k)$ and $j \in \mathcal{I}_V$, the assignment $\pi_{ij}^k$ is a function of the following form:

$$\pi_{ij}^k : \mathcal{X} \times \mathcal{Y}_k \rightarrow [0, 1],$$

(41)

where $\mathcal{X} \times \mathcal{Y}_k$ 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 $\mathcal{I}_T(t_k) = \emptyset$, i.e. there doesn’t exist any target in the mission space, there is no issue for cooperation and assignment. Therefore, let $t_k$ be a time instant at which $\mathcal{I}_T(t_k) \neq \emptyset$. 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.

$$\sum_{j \in \mathcal{I}_V} \pi_{ij}^k(x^k, y^k) = 1, \quad \forall i \in \mathcal{I}_V,$$

(42)

where $x^k$ and $y^k$ 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.

$$\sum_{j \in \mathcal{I}_V} \pi_{ij}^k(x^k, y^k) \leq 1, \quad \forall i \in \mathcal{I}_T(t_k).$$

(43)

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 rewards by acting generously. Therefore, it is required to over-assign the tasks to the vehicles, i.e.

$$\sum_{j \in \mathcal{I}_V} \pi_{ij}^k(x^k, y^k) \geq 1, \quad \forall i \in \mathcal{I}_T(t_k).$$

(44)

Remark 6: It can be shown that the inequalities in equations (43) and (44) turn to equalities when $|\mathcal{I}_V| = |\mathcal{I}_T(t_k)|$.

Equations (42), (43) and (44) introduce a set of constraints that should be satisfied by any desired assignment. More precisely, if one defines the set $\mathcal{P}^{n \times m}$ as

$$\mathcal{P}^{n \times m} = \{ \mathbf{\Pi} \in [0, 1]^{n \times m} : \mathbf{\Pi}^T \mathbf{1}_n = \mathbf{1}_m \},$$

(45)

for any $n, m \in \mathbb{N}$, then the assignment matrix $\mathbf{\Pi}^k(x^k, y^k)$ is required to belong to the set $\mathcal{P}_{\mathcal{I}_T(t_k), \mathcal{I}_V}$ which is defined as

$$\mathcal{P}_{\mathcal{I}_T(t_k), \mathcal{I}_V} = \{ \mathbf{\Pi} : \mathcal{X} \times \mathcal{Y}_k \rightarrow \mathbb{R}^{|\mathcal{I}_T(t_k)| \times |\mathcal{I}_V|} \}.$$  

(46)

D. 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 $u^k$.

Let $t_k$ be a time instant such that $\mathcal{I}_T(t_k) \neq \emptyset$ and $(u^k_j(\cdot))_{j \in \mathcal{I}_V}$ 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 $x^{k+1}$, is derived as in (22). Moreover, equation (24) provides the vector of predictions of position, velocity and acceleration of targets at $t_{k+1}$, the vector $\hat{y}^{k+1}$. Consider the estimation of the final maximum rewards introduced in Corollary 1, i.e. $\mathcal{R}_0 \rho_i(\tau_{ij}^k(u^k, t_k))$ 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 $u^k_j(\cdot)$ is applied for time interval $I_k = [t_k, t_{k+1})$, for any $i \in \mathcal{I}_T(t_k)$ and $j \in \mathcal{I}_V$. Also, consider the optimal assignment matrix for time instant $t_{k+1}$, denoted by $\hat{\mathbf{\Pi}}^{k+1}$, defined as the optimal assignment matrix. This matrix
is determined based on the state vector $x^{k+1}$ as well as the prediction vector $\hat{x}^{k+1}$ 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 $u^k$ and time instant $t_k$, i.e. $\hat{\Pi}^{k+1} = \hat{\Pi}^{k+1}(u^k, t_k)$. 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 $\mathcal{R}^{k+1}(u^k, t_k)$ and formulated as following

$$\mathcal{R}^{k+1}(u^k, t_k) = \sum_{i \in I} \sum_{j \in J} R_{ij}(\tau_{ij}(u^k, t_k)) x_{ij}(u^k, t_k), \quad (47)$$

where $\hat{R}_{ij}(u^k, t_k)$ is the entry of matrix $\hat{\Pi}^{k+1}(u^k, t_k)$ in row $i$ and column $j$, for any $i \in I$ and $j \in J$.

**Remark 7:** Note that $\mathcal{R}^{k+1}(u^k, t_k)$ is indeed an estimation performed at the current time, $t_k$, for the value of maximum total reward that the team expect at the next time instant $t_{k+1}$ to be capable of collecting by the end of mission.

Now, denote $P^k$ be the optimization problem for CRH controller corresponding to $k^{th}$ step. According to the discussion above, the following results:

$$p^k := \begin{cases} \max & \mathcal{R}^{k+1}(u^k, t_k) \\ \text{s.t.} & \hat{\Pi}(u^k, t_k) \in \mathcal{P}^k, u^k \in U^k, \end{cases} \quad (48)$$

where $\mathcal{P}^k$ and $U^k$ denote $\mathcal{P}_{I \times J}(t_k, I \times V)$ and $x_{j \in J} \mathcal{U}_k$, respectively.

**IV. Stationary Analysis of Cooperative Receding Horizon Controller**

The behavior of CRH controller which constructs the state trajectory of the system, depends on the parameters of the problem’s and the level of uncertainties in the environment. Given the parameters introduced in the problem formulation beside the time of arrivals and trajectories of targets, it is required to decide upon the value of planning horizons. In fact, the convergence of system is guaranteed only under special conditions, such as the proper choice of planning horizons. In order to demonstrate these conditions, the stationary analysis of the proposed receding horizon controller, for the case of a single vehicle is discussed in this section.

**A. Stationarity Property**

In the following the stationarity properties for the CRH controller are introduced.

**Definition 2:** The state trajectory $x(t) = (x(t))_{t \in I \times J}$ is called a strongly stationary trajectory if $\tau_t \in \mathcal{R}_{\geq 0}$, for any $i \in I$, i.e. each of the targets is hit by at least on of the vehicles in finite time. It is said that the CRH controller has a strong stationarity property if the state trajectory constructed is strongly stationary.

**Definition 3:** The state trajectory $x(t) = (x(t))_{t \in I \times J}$ is weakly stationary if $\tau_t \in \mathcal{R}_{\geq 0}$, for some $i \in I$, i.e. there is a target which is hit by at least on of the vehicles in finite time. The CRH controller introduced in (48) is said to have weak stationarity property if the resulted state trajectory is weakly stationary.

The strong stationarity is the desired property that the CRH controller is expected to have. However, roughly speaking, the weak stationarity of CRH controller guarantees its strong stationarity. Thus, it is enough to verify the weak stationarity property.

**B. A Technical Lemma**

Before proceeding to the stationarity analysis of CRH controller, see the following technical Lemma. This Lemma is a straight generalization of Lemma 1 in [12].

**Lemma 1:** Consider the vectors $p, q, r, v, a \in \mathbb{R}^d$ and the non-negative real numbers $t, m_2, m_3, \epsilon \in \mathbb{R}_{\geq 0}$ such that $\epsilon < m_2$. Let $\gamma : [0, \bar{t}] \rightarrow \mathbb{R}^d$ and $\hat{\gamma} : [0, \bar{t}] \rightarrow \mathbb{R}^d$ be three times continuously differentiable functions given by $\gamma(t) = r + vt + \frac{1}{2}at^2$ and $\hat{\gamma}(t) = r + vt + \frac{1}{2}a\bar{t}t^2$, where $|\gamma(t)| \leq m_2$, $|\hat{\gamma}(t)| \leq m_2 - \epsilon$ and $|\hat{\gamma}(t)| \leq m_3$, for any $t \in [0, \bar{t}]$, and also, $j : [0, \bar{t}] \rightarrow \mathbb{R}^d$ is a bounded three times continuously differentiable vector-valued function defined over the interval $[0, \bar{t}]$ with $\lim_{t \to 0} j(t) = 0$ and $\max_{t \in [0, \bar{t}]} ||j(t)|| \leq m_3$, for any $t \in [0, \bar{t}]$. Assume that $\gamma < \min\{z, c_1\}$ where $z$ is the smallest positive real scalar satisfying the equation $\frac{1}{2}(2m_2 - \epsilon)z^2 = ||(v - q)z + r - p||$. For any $t \in [0, \bar{t}]$, consider the following system of equations

$$\frac{1}{2}\dot{u}t^2 = -\frac{1}{2}\gamma(t)^2 + 2\gamma(t) - ut - q|t - t^2 - qt - p),$$

$$||u|| = m_2,$$

and also the system of equations

$$\frac{1}{2}\ddot{u}t^2 = \frac{1}{2}\hat{\gamma}(t)^2 + 2\hat{\gamma}(t) - ut - q|t - t^2 - qt - p),$$

$$||u|| = m_2.$$  

(50)

Then, for any $t \in [0, \bar{t}]$, each of the system of equations (49) and (50) has a solution in $\mathbb{R}_{\geq 0} \times \mathbb{R}^d$. Denoted by $(\tau(t), u(t))$ and $(\hat{\tau}(t), \hat{u}(t))$, respectively, as the solutions with smallest possible $\tau(t)$ and $\hat{\tau}(t)$. Also, for any $t \in [0, \tau(0])$, one has $\hat{\tau}(t) = \hat{\tau}(0) - t$ and $\hat{u}(t) = \hat{u}(0)$. Moreover, there exist continuous functions $\delta_\tau : [0, \bar{t}] \rightarrow \mathbb{R}$ and $\delta_u : [0, \bar{t}] \rightarrow \mathbb{R}$ such that $\tau = \hat{\tau} + \delta_\tau, u = \hat{u} + \delta_u$, and $\lim_{t \to 0} t^{-1}\delta_\tau(t) = \lim_{t \to 0} t^{-1}\delta_u(t) = 0$. Besides, for any $s \in \mathbb{R}_{\geq 0}$, there exist $t_s \in [0, \bar{t}]$ and $c_s \in (0, 1)$, such that if $||p - r|| > s$, one has that $\delta_\tau < (1 - c_s)t$, for any $t \in [0, t_s]$, and therefore $\tau(0) - \tau(t) \geq c_s t$.  

**Proof:** See Appendix.

**C. Stationary Analysis: Single Vehicle Case**

In the case that there is only a single vehicle in the mission space, one has $I \times V = \{1\}$. If $I \times R = \emptyset$, nothing remains to analysis. Therefore, let $I \times R \neq \emptyset$. Consider time instant $t_k$. If $I \times R(t_k) = \emptyset$, i.e. there doesn’t exist any target in the mission space, the vehicle is not required to make-decision and plan its path. So, let $t_k$ be a time instant at which $I \times R(t_k) \neq \emptyset$. For
brevity of notations, denote the assignments of the vehicles by \((\tau_{i}^k)_{i \in \mathcal{I}(t_k)}\), instead of \((\pi_{i1})_{i \in \mathcal{I}(t_k)}\). It can be easily verified that the constraints in (42), (44) and (43) reduce to the constraint
\[
\sum_{i \in \mathcal{I}(t_k)} \pi_i^k (u_i^k, y_i^k) = 1. \tag{51}
\]
Thus, the optimization problem of CRH controller, \(P_k\), becomes as follows
\[
\max \, \mathcal{R}(u_i^k, t_k) = \sum_{i \in \mathcal{I}(t_k)} \mathcal{R}_i \rho_i(\pi_i^k(u_i^k, t_k)) \tau_i^k(u_i^k, t_k), \tag{52}
\]
s.t. \(\pi_i^k(u_i^k, t_k) \in \Delta_{\mathcal{I}(t_k)}\), \(u_i^k \in \mathcal{U}_{I_k}\),
where \(\pi_i^k(u_i^k, t_k)\) denotes the vector \((\pi_{i1}^k(u_i^k, t_k))_{i \in \mathcal{I}(t_k)}\) and \(\Delta_m\) is the set \(\{a \in \mathbb{R}^m_0; a^\top 1_m = 1\}\), for any \(m \in \mathbb{N}\).

The feasible set in the optimization problem (52) is an infinite dimensional space. However the following Lemma introduces an optimization problem with finite dimensional feasible set where its solutions produce a solution for (52). This makes the problem significantly simpler to solve.

**Lemma 2:** Let the function \(\mathcal{R}(u_i^k, t_k)\) be defined as \(\mathcal{R}(u_i^k, t_k) = \sum_{i \in \mathcal{I}(t_k)} \mathcal{R}_i \rho_i(\pi_i^k(u_i^k, t_k)) \tau_i^k(u_i^k, t_k)\). Also, define the function \(f\) as \(f(\tau, p, q, r, v, a) = 1/2 \tau^2 + (v - q) \tau + (r - p)\). Consider the following optimization problem
\[
\max \, \mathcal{R}(\tau, \pi) = \sum_{i \in \mathcal{I}(t_k)} \mathcal{R}_i \rho_i(\pi_i^k(u_i^k, t_k)) \tau_i^k(u_i^k, t_k), \tag{53}
\]
s.t. \(\|f(\tau, p, q, r, v, u_i^k, a_i^k)\| \leq \frac{1}{2} \tau_i^2 u_{\max}, \forall i \in \mathcal{I}(t_k)\),
\(\tau \in \mathbb{R}^{\mathcal{I}(t_k)}_0\),
\(\pi \in \Delta_{\mathcal{I}(t_k)}\).

Then, (53) has a solution like \((\tau^*, \pi^*)\) such that there exists \(i^k \in \mathcal{I}(t_k)\) where all entries of \(\pi^*\) are zero except \(\pi_{i^k}^*\) which is one. Moreover, there exists a vector \(u^k \in \mathbb{R}^d\) such that \(\|u^k\| = u_{\max}\) and \(f(\tau^*, p^k, q^k, r^k, v^k, u^k) = \frac{1}{2} (\tau_{i^k}^*)^2 u^k\).

Also, if \(t_{k+1} - t_k < i^k\) where \(i^k\) is the index introduced for \(i\) in Lemma 1 with the vectors \(p^k, q^k, r^k, v^k, u^k\) and the scalars \(u_{\max}, a, \gamma\), then the control input \(u(t) \equiv u^k\) is a solution for the optimization problem (52).

**Proof:** Let \(\mathcal{F}\) be the subset of \(\mathbb{R}^{\mathcal{I}(t_k)}_0\) which is defined as
\[
\mathcal{F} = \{\tau \in \mathbb{R}^{\mathcal{I}(t_k)}_0; ||f(\tau, p, q, r, v, u_i, a_i)|| \leq \frac{1}{2} \tau^2 u_{\max}, \forall i \in \mathcal{I}(t_k)\}, \tag{54}
\]
and also, denote by \(\mathcal{F}\) the feasible set for the problem (53) which is \(\mathcal{F} \times \Delta_{\mathcal{I}(t_k)}\). Since \(f\) and the norm are continuous functions, set \(\mathcal{F}\) is closed. Subsequently, as set \(\Delta_{\mathcal{I}(t_k)}\) is compact, \(\mathcal{F}\) is a closed set. Now, let \(i\) be in \(\mathcal{I}(t_k)\). For any \(\tau_i \in \mathbb{R}_{\geq 0}\), from triangle inequality, that
\[
||f(\tau_i, p^k, q^k, r^k, v^k, a_i^k)|| \leq \frac{1}{2} ||a_i^k||^2 + ||v^k - q^k|| \tau_i + ||r^k - p^k||. \]

Since \(||a_i^k|| < u_{\max}\) for large enough \(\tau_i\), one has
\[
\frac{1}{2} ||a_i^k||^2 + ||v^k - q^k|| \tau_i + ||r^k - p^k|| \leq \frac{1}{2} u_{\max} \tau_i^2.
\]
and therefore the following inequality holds:
\[
||f(\tau_i, p^k, q^k, r^k, v^k, a_i^k)|| \leq \frac{1}{2} u_{\max} \tau_i^2.
\]
This shows that for large enough \(b \in \mathbb{R}_{\geq 0}\), set \(\mathcal{F}_b\) defined as
\[
\mathcal{F}_b = \mathcal{F} \cap [0, b]^{\mathcal{I}(t_k)}
\]
is non-empty. Moreover, it is closed and bounded. Therefore, \(\mathcal{F}_b\) which is defined as \(\mathcal{F}_b \times \Delta_{\mathcal{I}(t_k)}\) is a non-empty compact set. From continuity of \(\mathcal{R}^{k+1}\), one has that the set \(\mathcal{M}_b\) defined as followings
\[
\mathcal{M}_b = \arg\max_{(\tau, \pi) \in \mathcal{F}_b} \mathcal{R}^{k+1}(\tau, \pi).
\]
Also, since \(\tau'' \leq \tau'\) and each of the discount functions is a decreasing function, one has
\[
\mathcal{R}^{k+1}(\tau', \pi') \leq \mathcal{R}^{k+1}(\tau'', \pi''),
\]
and therefore, one can verify from (55) and (56) that the following inequality holds
\[
\mathcal{R}^{k+1}(\tau', \pi') \leq \mathcal{R}^{k+1}(\tau, \pi).
\]
This shows that the set of solutions of (53) is a non-empty set. Accordingly, one can define set \(\mathcal{M}^{k+1}\) as a subset of \(\mathcal{F}^{k+1}\) with following property:
\[
\mathcal{M}^{k+1} = \arg\max_{(\tau, \pi) \in \mathcal{F}^{k+1}} \mathcal{R}^{k+1}(\tau, \pi).
\]
Set \((\tau, \pi)\) as an arbitrary element of \(\mathcal{M}^{k+1}\). Let \(i^k\) be defined as
\[
i^k = \min_{i \in \mathcal{I}(t_k)} \mathcal{R}_i \rho_i(\pi_i).
\]
Accordingly, let \(\pi^*\) be a vector in \(\Delta_{\mathcal{I}(t_k)}\) such that each of its entries is zero except the entry corresponding to \(i^k\) which is 1. Also, set vector \(\pi^*\) equal with \(\pi\). It can be verified that
\[
\mathcal{R}_i \rho_i(\pi_i) \leq \mathcal{R}_i \rho_i(\pi_i^k) = \mathcal{R}^{k+1}(\tau^*, \pi^*).
\]
Since \(\pi^{1}_{i \in \mathcal{I}(t_k)} = 1.1\), it is concluded that
\[
\mathcal{R}^{k+1}(\tau, \pi) = \sum_{i \in \mathcal{I}(t_k)} \mathcal{R}_i \rho_i(\pi_i) \tag{61}
\]
\[
\leq \mathcal{R}_i \rho_i(\pi_i^k) = \mathcal{R}^{k+1}(\tau^*, \pi^*).\]
From \((\tau, \pi) \in \mathcal{M}^{k+1}\) and \((\tau^*, \pi^*) \in \mathcal{F}^{k+1}\), one has
\[
\mathcal{R}^{k+1}(\tau^*, \pi^*) \leq \mathcal{R}^{k+1}(\tau, \pi) \tag{62}
\]
This shows that \((\tau^*, \pi^*) \in \mathcal{M}^{k+1}\). If the following equality holds
\[
||f(\tau_{i^k}, p^k, q^k, r^k, v^k, a_i^k)|| = \frac{1}{2} (\tau_{i^k})^2 u_{\max}, \tag{63}
\]
then one can set \(u_k\) as
\[
u_k = \frac{2}{(\tau_{i^k})^2} f(\tau_{i^k}, p^k, q^k, r^k, v^k, a_i^k), \tag{64}
\]
which is a vector with equal to \( u_{\max} \) and satisfies the desired properties. If the equality (63) does not hold, one must have

\[
\|f(\tau_k, p^k, q^k, r^k, s^k, a^k)\| < \frac{1}{2}(\tau_k^*)^2 u_{\max}. \tag{65}
\]

Now, define the continuous function \( g : [0, \tau_k^*] \to \mathbb{R} \) as

\[
g(\tau) = \|f(\tau, p^k, q^k, r^k, s^k, a^k)\| - \frac{1}{2}\tau^2 u_{\max}. \tag{66}
\]

Since \( g(0) > 0 \) and \( g(\tau_k^*) < 0 \), there exists \( \bar{\tau} \in (0, \tau_k^*) \) such that

\[
g(\bar{\tau}) = 0, \quad i.e.
\]

\[
\|f(\bar{\tau}, p^k, q^k, r^k, s^k, a^k)\| = \frac{1}{2}\bar{\tau}^2 u_{\max}. \tag{67}
\]

Note that as \( \bar{\tau} < \tau_k^* \) and discount function \( \rho_t \) is a decreasing function, it is concluded that \( \rho_t(\tau_k^*) \leq \rho_t(\bar{\tau}) \) and subsequently,

\[
\tilde{\Omega}^{k+1}(\tau^*, \pi^*) = R_k \rho_t(\tau_k^*) \leq R_k \rho_t(\bar{\tau}) = \tilde{\Omega}^{k+1}(\tau^*, \pi^*), \tag{68}
\]

where \( \pi \) is the vector where each of its entries is equal to the respective entry in \( \pi^* \) except the entry corresponding to \( i^k \) which is \( \bar{\tau} \). Substituting \( \pi^* \) with \( \pi \) and obtaining \( u_k^* \) from equation (64), vector \( u_k^* \) with desired properties results.

Let \( u^*(t) \in \mathcal{U}_k \) be a solution for (52) and \( \pi^* \) be the resulting assignment vector. Define \( i^* = \arg\max_{i \in \mathcal{I}_T(t_k)} \Omega_i \) as following

\[
i^* = \min \arg\max_{i \in \mathcal{I}_T(t_k)} \mathcal{R}_i \rho_i(\tau_i(u^*(\cdot), t_k)), \tag{69}
\]

where \( \tau_i(u^*(\cdot), t_k) \) is the time estimate when the vehicle can reach target \( i \), based on the information given at time instant \( t_k \) and being the control input \( u^*(\cdot) \) applied by the vehicle. Similar to the discussion given above, one can show that

\[
\Omega^{k+1}(u^*, t_k) = \mathcal{R}_i \rho_i(\tau_i(u^*(\cdot), t_k)), \tag{70}
\]

and also, \( u^*(t) \) is a solution for the following optimization problem:

\[
\max \quad \mathcal{R}_i \rho_i(\tau_i(u^*(\cdot), t_k)) \\
\text{s.t.} \quad u^*(\cdot) \in \mathcal{U}_k, \tag{71}
\]

Since, otherwise there exist \( u(\cdot) \in \mathcal{U}_k \) such that

\[
\mathcal{R}_i \rho_i(\tau_i(u^*(\cdot), t_k)) < \mathcal{R}_i \rho_i(\tau_i(u(\cdot), t_k)) \leq \Omega^{k+1}(u(\cdot), t_k), \tag{72}
\]

and therefore,

\[
\Omega^{k+1}(u^*, t_k) = \mathcal{R}_i \rho_i(\tau_i(u^*(\cdot), t_k)) < \mathcal{R}_i \rho_i(\tau_i(u(\cdot), t_k)) \leq \Omega^{k+1}(u(\cdot), t_k), \tag{73}
\]

which contradicts the fact that \( u^*(\cdot) \) is a solution for (52).

Since \( \rho_t \) is a strictly decreasing function, \( u^*(\cdot) \) is a solution for the following optimization problem:

\[
\min \quad \tau_i(u(\cdot), t_k) \\
\text{s.t.} \quad u(\cdot) \in \mathcal{U}_k, \tag{74}
\]

Since \( t_{k+1} - t_k \leq i^k \), then the vehicle does not visit target \( i^* \) prior to time \( t_{k+1} \). Hence, one has \( \tau_i(u^*(\cdot), t_k) \geq t_{k+1} \). Let \( x(t_{k+1}) = [p^T(t_{k+1}), q^T(t_{k+1})]^T \) be the state vector of vehicle after applying control input \( u^* \) for time period \( I_k \). Accordingly, \( \tau_i(u^*(\cdot), t_k) - t_{k+1} \) is the estimated minimum time for reaching to target \( i^* \) starting from \( x(t_{k+1}) \). From Bellman’s principle of optimality, the solution of (74) gives the smallest estimated minimum reaching time starting from \( x(t_k) \). Therefore, if \( u^* \) represents the optimal control input for time interval after \( t_{k+1} \) under which the vehicle reaches the anticipated position of target \( i^* \) in \( \tau_i(u^*(\cdot), t_k) - t_{k+1} \) unit of time, then the control input defined as

\[
u = u^1 I_k + u^{**} 1_{g, t \geq t_{k+1}} \tag{75}
\]

is a solution for the following minimum time optimal control problem

\[
\min \quad \mathcal{J}_i (u) = \int_{t}^{T} 1 \, ds \\
\text{s.t.} \quad \frac{d}{dt} x = A x + B u, \\
\quad w(0) = x(t_k), \\
\quad \Psi_{i+1}^{k+1}(\tau, w(\tau)) = 0, \\
\quad u \in \mathcal{U}(u_{\max}), \tag{76}
\]

where function \( \Psi_{i+1}^{k+1} : \mathbb{R}_+ \times \mathbb{R}^{2d} \to \mathbb{R}^d \) is defined as

\[
\Psi_{i+1}^{k+1}(t, x) = |D_i 0 d| x - \frac{1}{2} i^{k+1} t^2 - \bar{v}_i^{k+1} t - \bar{v}_i^{k+1}. \tag{77}
\]

for any \( i \in \mathcal{T}(t_k) \). Based on a discussion similar to the one given in the proof of Theorem 1, one can show that the optimal control input \( u_i(\cdot) \) is a constant function where its value is given by (64). This concludes the proof.

For the sake of simplicity, take the discount function as in (16). Before proceeding verifying the stationarity of the proposed receding horizon scheme in (52), consider the following Lemma.

**Lemma 3**: Let \( n \) be a positive integer, \( \Omega \) be a subset of \( \mathbb{R}^n \), \( \{c_i\}_{i \in N_n} \) and \( \{\lambda_i\}_{i \in N_n} \) be positive real scalars. Define the function \( g_1, g_2 : \Omega \times \Delta_n \to \mathbb{R} \) as \( g_1(x, a) = \sum_{i=1}^{n} c_i a_i e^{-\lambda_i x_i} \), and \( g_2(x, a) = \sum_{i=1}^{n} a_i (\lambda_i x_i - \ln c_i) \). Then, \( (x^*, a^*) \) is a maxmizer of \( g_1 \) over \( \Omega \times \Delta_n \) if and only if it is a minimizer of \( g_1 \) over the same set. Moreover, for any such \( (x^*, a^*) \), there exists some \( i \in N_n \) such that \( g_1(x^*, a^*) = g_1(x^*, a_i) \), \( a_i^* = g_2(x^*, a_i) \).

**Proof**: Let \( (x^*, a^*) \) be a maximizer of \( g_1 \) over \( \Omega \times \Delta_n \) and \( i^* \in N_n \) be such that \( c_i e^{-\lambda_i x_i} \leq c_i e^{-\lambda_i x_i^*}, \) for any \( i \in N_n \). Since, \( 1^T a^* = 1 + a^* \geq 0 \), it can be easily verified that

\[
\begin{align*}
g_1(x^*, a^*) &= \sum_{i=1}^{n} c_i a_i^* e^{-\lambda_i x_i^*} \\&\leq c_i e^{-\lambda_i x_i} - c_i e^{-\lambda_i x_i^*} \\&= g_1(x^*, a_i),
\end{align*}
\]

where \( c_i^* \) is a vector where its \( (i^*)^n \) entry is 1 and all of its other entries are zero. Also, from \( (x^*, a_i^*) \in \Omega \times \Delta_n \), it is deduced that \( g_1(x^*, a^*) \geq g_1(x^*, e_i) \), and subsequently, \( g_1(x^*, a_i^*) = g_1(x^*, e_i^*) \), i.e. \( (x^*, a_i^*) \) is a maximizer of \( g_1 \) over \( \Omega \times \Delta_n \). Note that for any \( i \in N_n \) such that \( c_i e^{-\lambda_i x_i} < c_i e^{-\lambda_i x_i^*} \), it is required that \( a_i^* = 0 \), since otherwise one has

\[
g_1(x^*, a_i^*) < g_1(x^*, e_i),
\]

which contradict the assumption that \( (x^*, a^*) \) is a maximizer of \( g_1 \) over \( \Omega \times \Delta_n \).

Now, let \( x \) be an arbitrary element of \( \Omega \). Then, for any \( i \in N_n \), one has \( g_1(x, e_i) \leq g_1(x^*, e_i^*) \), i.e. \( c_i e^{-\lambda_i x_i} \leq c_i e^{-\lambda_i x_i^*} \).
Considering the introduced performance index, it is deduced that \( \lambda_i x_i - \ln c_i \geq \lambda_i x_i^* - \ln c_i^* \), for any \( i \in \mathbb{N} \). Therefore, for any \( a \in \Delta_n \), one has
\[
g_2(x, a) = \sum_{i=1}^{n} a_i (\lambda_i x_i - \ln c_i) \geq \lambda_i x_i^* - \ln c_i^* = g_2(x^*, e_i^*),
\]
and therefore, \((x^*, e_i^*)\) is a minimizer of \( g_2 \) over \( \Omega \times \Delta_n \). Since, \( a_i^* = 0 \), for any \( i \in \mathbb{N} \), such that \( \lambda_i x_i - \ln c_i \neq \lambda_i x_i^* - \ln c_i^* \), it can be easily verified that \( g_2(x^*, e_i^*) = g_2(x^*, a^*) \) and thus, \((x^*, a^*)\) is a minimizer of \( g_2 \) over \( \Omega \times \Delta_n \).

The other direction of theorem can be shown using a similar discussion and the proof is concluded.

Before proceeding to introducing the main theorem and its proof, for the sake of simplicity, let assume that for some positive scalar \( c \) one has \( \mathcal{R} \mathcal{E}^\gamma = c \), for any \( i \in \mathcal{I}_t \), i.e. the targets are assumed to have same levels of priority.

**Theorem 2:** Consider the receding horizon problem presented in (52). Assume that \( \bar{a} < u_{\max} \). Let also \( H_k < \bar{a} \) for any \( k \) where \( H_k \) is the bound introduced for \( I \) in Lemma 1 with the vectors \( p^k, q^k, \rho^k, v^k, a^k \) and the scalars \( u_{\max}, \bar{a}, \bar{f} \). Then for any initial choice of \( x \) and \( \pi \) and satisfying (51), the receding horizon algorithm is finite-time convergent, i.e., the vehicles reach the targets in finite time.

**Proof:** From Lemma 2, it is concluded that at each iteration of decision-making, one can obtain the solution by solving the equivalent optimization problem (53). Moreover, Lemma (3) says that the optimization problem (53) is equivalent with following optimization problem:
\[
\min_{\mathcal{I}_{\mathcal{T}}(t_k)} \mathcal{I}_k \pi = \sum_{i \in \mathcal{I}_{\mathcal{T}}(t_k)} (\gamma_t \tau_i - \ln q_i - \gamma_t \tau_i^*) \pi_i
\]
\[
s.t. \quad \|f(\tau_i, p^{k^*}, q^{k^*}, \rho^{k^*}, v^{k^*}, a^{k^*})\| \leq \sqrt{2} u_{\max}, \forall i \in \mathcal{I}_{\mathcal{T}}(t_k)
\]
\[
\tau \in \mathbb{R}_{\geq 0}^{\mathcal{I}_{\mathcal{T}}(t_k)}, \quad \pi \in \Delta(\mathcal{I}_{\mathcal{T}}(t_k)).
\]

Since, for any \( \pi \in \Delta(\mathcal{I}_{\mathcal{T}}(t_k)) \), one has \( \pi^T 1_{\mathcal{I}_{\mathcal{T}}(t_k)} = 1 \), it is concluded that
\[
\mathcal{I}_k \pi = \sum_{i \in \mathcal{I}_{\mathcal{T}}(t_k)} \gamma_t \tau_i \pi_i
\]
\[
= \left( \sum_{i \in \mathcal{I}_{\mathcal{T}}(t_k)} \gamma_t \tau_i \pi_i \right) - \ln c_i
\]
\[
(79)
\]

For any \( k \in \mathbb{N} \), define the performance index \( \mathcal{P}_k \) as a function \( \mathcal{P}_k : [\mathcal{I}_{\mathcal{T}}(t_k)] \times [\mathcal{I}_{\mathcal{T}}(t_k)] \to \mathbb{R} 
\)
\[
\mathcal{P}_k(\tau, \pi) = \sum_{i \in \mathcal{I}_{\mathcal{T}}(t_k)} \gamma_t \tau_i \pi_i
\]
\[
(80)
\]

Considering the introduced performance index, the optimization problem (78) is simplified to the following problem:
\[
\min_{\mathcal{I}_{\mathcal{T}}(t_k)} \mathcal{P}_k(\tau, \pi)
\]
\[
s.t. \quad \|f(\tau_i, p^{k^*}, q^{k^*}, \rho^{k^*}, v^{k^*}, a^{k^*})\| \leq \sqrt{2} u_{\max}, \forall i \in \mathcal{I}_{\mathcal{T}}(t_k)
\]
\[
\tau \in \mathbb{R}_{\geq 0}^{\mathcal{I}_{\mathcal{T}}(t_k)}, \quad \pi \in \Delta(\mathcal{I}_{\mathcal{T}}(t_k)).
\]

For any \( k \in \mathbb{N} \), let \((\tau^{k^*}, \pi^{k^*})\) be the solution of (81), \( u^{k^*} \) be the corresponding control input and \( i^{k^*} \) be the index introduced in Lemma 3.

Let assume that the vehicle never visits any of the targets. More precisely, for any \( t \) and any \( i \in \mathcal{I}_{\mathcal{T}}(t) \), one has \( \|p(t) - r_i(t)\| > d_i \). Therefore, for any \( t_k \) and any \( i \in \mathcal{I}_{\mathcal{T}}(t_k) \), one must have that \( \|p(t_k) - r_i(t_k)\| > s = \min_{i \in \mathcal{I}_{\mathcal{T}}} d_i \). Now, let \( k \in \mathbb{N} \). For \((\tau^{k^*}, \pi^{k^*})\), one has
\[
\mathcal{P}_k(\tau^{k^*}, \pi^{k^*}) = \mathcal{P}_k(\tau^{k^*}, e_i^{k^*}).
\]
\[
(82)
\]

Therefore,
\[
\mathcal{P}_k(\tau^{k^*}, \pi^{k^*}) = \mathcal{P}_k(\tau^{k^*}, e_i^{k^*})
\]
\[
(83)
\]

From Lemma 1, one has
\[
\tau^{k^*} - \tau^{k^*+1} \geq (1 - c_s)(t_k+1 - t_k),
\]
\[
(84)
\]

where \( \tau^{k^*+1} \) is the \((i^{k^*})\)th entry of \( \tau^{k^*+1} \) and \( \tau^{k^*+1} \) is defined as the solution of following optimization problem:
\[
\min_{\mathcal{I}_{\mathcal{T}}(t_k)} \mathcal{P}_k^{k+1}(\tau, e_i^{k^*})
\]
\[
s.t. \quad \tau \in \mathbb{R}_{\geq 0}^{\mathcal{I}_{\mathcal{T}}(t_k)},
\]
\[
\|f(\tau_i, p^{k^*+1}, q^{k^*+1}, \rho^{k^*+1}, v^{k^*+1}, a^{k^*+1})\| \leq \sqrt{2} u_{\max}, \forall i \in \mathcal{I}_{\mathcal{T}}(t_k+1)
\]
\[
(85)
\]

Therefore, it is deduced that
\[
\mathcal{P}_k^{k+1}(\tau, e_i^{k^*}) \geq \gamma_i \tau_i^{k^*} + \gamma_i (1 - c_s)(t_k+1 - t_k),
\]
\[
(86)
\]

or equivalently,
\[
\mathcal{P}_k^{k+1}(\tau, e_i^{k^*}) \geq \gamma_i \tau_i^{k^*+1} + \gamma_i (1 - c_s)(t_k+1 - t_k),
\]
\[
(87)
\]

Since, one has \( \mathcal{P}_k^{k+1}(\tau^{k^*+1}, e_i^{k^*}) \geq \mathcal{P}_k^{k+1}(\tau^{k^*+1}, \pi^{k^*+1}), \) it is concluded that
\[
\mathcal{P}_k^{k+1}(\tau^{k^*+1}, \pi^{k^*+1}) \geq \gamma_i \tau_i^{k^*} + \gamma_i (1 - c_s)(t_k+1 - t_k),
\]
\[
(88)
\]

where \( \gamma_i \) is defined as \( \gamma_i \in \mathcal{I}_{\mathcal{T}} \). By induction, for any \( m \in \mathbb{N} \), one can show that
\[
\mathcal{P}_k^{m}(1 - c_s)(t_m - t_{m-1}),
\]
\[
(89)
\]

and also
\[
\mathcal{P}_k^{m}(1 - c_s)(t_m - t_{m-1}) \geq \mathcal{P}_k^{m}(\tau^{m}, \pi^{m}).
\]
\[
(90)
\]

Therefore, it can be seen that
\[
\lim_{m \to \infty} \mathcal{P}_k^{m}(\tau^{m}, \pi^{m}) = -\infty
\]
\[
(91)
\]

For any \( m \in \mathbb{N} \), one has that \( \tau^{m} \) and \( \pi^{m} \) are vectors with non-negative entries. Therefore, the performance index takes non-negative values, i.e.
\[
\mathcal{P}_k^{m}(\tau^{m}, \pi^{m}) = \sum_{i \in \mathcal{I}_{\mathcal{T}}(t_m)} \gamma_i \tau_i^{m} \pi_i^{m} \geq 0,
\]
\[
(92)
\]

which contradicts (91). Therefore, there must be a target which is visited in some time instant \( t_k \). Reconsidering the problem with remaining targets, the same result holds and there must be another target which is visited in some time instant \( t_{k+2} \), etc. Hence, as there are finite number of targets, all of them are visited in finite number of steps.

\[\square\]
V. SIMULATION RESULTS

In this section, three simulation scenarios are 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. These scenarios are designed to show the effectiveness and flexibility of the proposed method in meeting various dynamic decision criteria solely by modifying the reward functions to the most appropriate. In all of these scenarios, the square \( \mathcal{M} = [-200, 200] \times [-200, 200] \) in the flat plane 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 unknown trajectory and arrival time; however, at each time instant \( t \) the information vector of the targets \( y(t) \) (see (12)) are updated. The maximum acceleration of vehicles and targets are bounded by \( u_{\text{max}} = 2m/s^2 \) and \( a = 1m/s \), respectively. It is assumed that the targets always satisfy Assumptions 2 and 3. Initially two targets are also present in the mission space along with the vehicles, and the remaining two targets arrive sequentially at \( \{\hat{\tau}_i\}_{i=3}^4 = \{2, 4\} \). Finally the initial reward of each target is the same and equal to \( \{ \rho_i = 1 \}_{i=1}^4 \). Note that all of the aforementioned initializations and conditions are maintained in the same in all three scenarios for the sake of comparison.

Example 1 (Exponential reward function): In this example, it is assumed that the targets have same reward function which are in the form of equation (16) with reward discount rate parameter \( \{\gamma_i = 1\}_{i=1}^4 \). Fig. 1 depicts the position of the 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: \( T_s = 0.05 \)). Fig. 2 shows the vehicle-target assignments through the simulation time period. Note that since the result of algorithm in terms of \( \{\pi_{i,j}^{k}\} \) is either 0 or 1, for each vehicle \( j \), only one target is visited and the value of remaining assignments is 0. In other words the vehicles can be assigned to \( |I_T(t_k)\} \) targets at each time step \( k \). The total expected reward, (see (47)), the available total reward \( \sum_{i \in I_T} \rho_i(t) \) (see (16)) and the total collected reward are shown in Fig. 3, at each step of time. The result of this example shows that all of the targets are visited in finite time which are as follows: \( \{\hat{\tau}_i\}_{i=1}^4 = \{1.05, 2.85, 4.05, 6.05\} \). Moreover, the total reward of 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 \( t = 2 \).

Example 2 (Hard deadline reward function): The reward function in this example is in the form of hard deadline (see equation (17)) with \( \{t_{f,j} = 8s\}_{j=1}^4 \). Figs. 4 and 5 show the position of the vehicles and targets and the vehicle-target assignments through the simulation time period, respectively. The total expected reward, (see (47)), the total available reward \( \sum_{i \in I_T} \rho_i(t) \) (see (17)) and the total collected reward are shown in Fig. 6, at each step of time. Note that the total available reward is the sum of the targets’ linear time-decreasing reward functions and if a target can not be visited before the
corresponding deadline, the algorithm will abandon it. Figs. 4, 5 and 6 show that all of the targets are visited in finite time which are as follows: \( \{\tilde{\tau}_i\}_{i=1}^4 = \{1.5, 2.95, 4.2, 8.9\} \). Moreover, the total reward of 2.181 is collected in this mission.
Example 3 (Soft deadline reward function): In this example reward functions with hard deadlines (see equation 17) \( \{t_i^f = 8s\}_{i=1} \) are considered for the first three targets while for the fourth target a reward function with an a priori unknown soft deadline is utilized as follows:

\[
\rho_4(t) = \begin{cases} 
\max\{1 - \frac{t - \hat{r}_4}{\hat{r}_4 - \tau_4}, 0\}, & \text{if } t \in [\hat{r}_4, D_4), \\
\max\{1 - \frac{t - \hat{r}_4}{\tau_4 - \hat{r}_4}, 0\}, & \text{if } t \in [D_4, \hat{r}_4), 
\end{cases}
\]

where \( D_4 = 2.5s \) is the time when the deadline of reward function change from \( t_4^f = 8s \) to \( t_4^f = 50s \). Figs. 4 and 5 show the position of the vehicles and targets and the vehicle-target assignments through the simulation time period, respectively. The total expected reward, (see (47)), the total available reward \( \sum_{i \in \mathcal{I}_T} \rho_i(t) \) (see (18)) and the total collected reward are shown in Fig. 9, at each step of time. Figs. 4, 5 and 6 show that all of the targets are visited in finite time which are as follows: \( \{\hat{r}_i\}_{i=1}^4 = \{2, 2.95, 4.85, 8.45\} \). Moreover, the total reward of 2.865 is collected in this mission. Comparing Figs. 5 and 8 divulges a shift in the decision making process of the algorithm when a change occurs in the reward function of target 4. The decreasing rate of the reward function of target 4 with respect to time become smaller, therefore, vehicle 2 looses interest from target 4 and aims to target 3. Roughly speaking, this decision altering phenomenon is happening since maximizing collecting reward will prioritizes targets with larger reward decreasing rates. A comparison between Figs. 6 and 9 will also shed some light on this phenomenon. Note that in examples 2 and 3 where the reward of targets are zero after some deadline, if any target is not reachable before deadlines by any vehicle, they will be abandoned. On the contrary, the exponentially decreasing reward in 3 will always be nonzero, which means that the targets are always considered in the dynamic decision making process.

With the help of the aforementioned simulation study, one can see the merits and efficiency of the proposed CRH controller with expected reward maximization scheme in generating the optimal assignment \( \Pi^k \) and optimal control input \( u^k \). More specifically, these examples show that the decreasing rate of reward function and the deadlines (if any) are critical in the proposed dynamic decision making process since the problem is formulated as a maximum reward collecting problem.

**VI. CONCLUSIONS**

In this work, the cooperative multi-target interception problem in uncertain environments with double-integrators vehicles is investigated. The uncertainty of the environment is derived from the fact that these targets have unknown arrival times and trajectories. Allocating a reward function to each target, the problem is reformulated as an optimization problem which maximize the expected reward collectible from the set of available targets in the mission space. The inertia of the double integrator vehicles and the uncertainty of the environment make the one-shot optimization impossible. Therefore a cooperative receding horizon controller is utilized. The proposed controller sequentially solves the optimization problem by estimating the collectible reward over a planning horizon and executes the control for an action horizon. The trajectories of the vehicles are shown to be stationary, and the effectiveness and advantages of the proposed controller was demonstrated via a simulation study.

**APPENDIX**

Define function \( \Psi : \mathbb{R} \times \mathbb{R}^{2d} \times \mathbb{R}^{2d} \to \mathbb{R}^d \) as

\[
\Psi(t, w^{(1)}, w^{(2)}) = [L_t, 0_d]w^{(1)} - [L_t, 0_d]w^{(2)}. 
\]

Consider the following optimal control problem:

\[
\min \int_{t}^{\tau} \left( u^{(1)}, u^{(2)} \right) = \int_{0}^{\tau} 1 \, ds \\
\text{s.t.} \quad \frac{d}{dt}w^{(1)} = A_dw^{(1)} + B_du^{(1)} \\
\frac{d}{dt}w^{(2)} = A_dw^{(2)} + B_du^{(2)}, \\
\frac{d}{dt}w^{(1)}(0) = [p^T, q^T]^T, \quad \frac{d}{dt}w^{(2)}(0) = [p^T, q^T]^T, \quad \|w^{(1)}\|_\infty \leq 2m_2, \quad \|w^{(2)}\|_\infty \leq 2m_2 - \epsilon. \\
\Psi(t, w^{(1)}(\tau), w^{(2)}(\tau)) = 0,
\]

and define the corresponding Hamiltonian function [16]

\[
H : \mathbb{R}_{\geq 0} \times \mathbb{R}^{2d} \times \mathbb{R}^{2d} \times \mathbb{R}^{2d} \times \mathbb{R}^{2d} \times \mathbb{R}^d \to \mathbb{R}
\]

with

\[
H(\lambda_0, \lambda^1, \lambda^2, w^{(1)}, w^{(2)}, u^{(1)}, u^{(2)}) = \lambda_0 + \lambda^1 \left[ A_dw^{(1)} + B_du^{(1)} \right] + \lambda^2 \left[ A_dw^{(2)} + B_du^{(2)} \right].
\]

The Pontryagin maximum principle [16] says that if \( u = (u^{(1)}, u^{(2)}) \) be the optimal solution of (95) and \( w = (w^{(1)}, w^{(2)}) \) be the resulting optimal state trajectory, then there exist a non-negative real scalar \( \lambda_0 \in \mathbb{R}_{\geq 0} \) and costate trajectories \( \lambda = (\lambda^1, \lambda^2) : [0, \tau] \to \mathbb{R}^{2d} \times \mathbb{R}^{2d} \) satisfying the following conditions:

i. One has \( (\lambda_0, \lambda^1(\tau), \lambda^2(\tau)) \neq 0 \), for any \( t \in [0, \tau] \).

ii. The costate trajectories satisfy the following differential equations

\[
\frac{d}{dt}\lambda^1 = -A_d^T\lambda^1, \quad \frac{d}{dt}\lambda^2 = -A_d^T\lambda^2,
\]

iii. The following equation holds

\[
H(\lambda_0, \lambda, w, u) = \min_{(\nu^1, \nu^2) \in \mathcal{G}_{m_2, m_2 - \epsilon}} H(\lambda_0, \lambda, w, \nu^1, \nu^2),
\]

for any \( t \in [0, \tau] \), where \( \mathcal{G}_{m_2, m_2 - \epsilon} \subset \mathbb{R}^{2d} \times \mathbb{R}^{2d} \) is defined as \( \mathcal{G}_{m_2, m_2 - \epsilon} \). Moreover, \( H(\lambda_0, \lambda, w, u) \) is constant.

iv. For some vector \( \nu \in \mathbb{R}^{2d} \), one has

\[
H = -\left( D_w \Psi \right) \tau T \nu,
\]

and

\[
\lambda(\tau) = (D_w \Psi) \tau T \nu.
\]

Let \( \bar{w}^{(1)} = [p^{(1)}T, q^{(1)}T]^T \), \( \bar{w}^{(2)} = [p^{(2)}T, q^{(2)}T]^T \), \( \lambda^1 = [\lambda_p^{(1)}T, \lambda_q^{(1)}T]^T \) and \( \lambda^2 = [\lambda_p^{(2)}T, \lambda_q^{(2)}T]^T \). Then, from equation (98) one can see that

\[
\begin{cases}
\frac{d}{dt}\lambda_p^{(1)} = 0_d, \\
\frac{d}{dt}\lambda_q^{(1)} = -\lambda_p^{(1)}, \\
\frac{d}{dt}\lambda_p^{(2)} = 0_d, \\
\frac{d}{dt}\lambda_q^{(2)} = -\lambda_p^{(2)},
\end{cases}
\]

where \( \lambda_p \) and \( \lambda_q \) are the costate variables associated with the position and velocity, respectively.
and therefore, $\lambda_p^{(1)}$ and $\lambda_p^{(2)}$ are constant. Also, $\lambda_q^{(1)}(t) = \lambda_q^{(1)}(0) - \lambda_q^{(2)} t$ and $\lambda_q^{(2)}(t) = \lambda_q^{(2)}(0) - \lambda_q^{(1)} t$, for any $t \in [0, \tau]$. Since one has

$$D_w \Psi(t, w) = [I_d, 0_d, -I_d, 0_d]^T, \quad (103)$$

it is deduced from equation (101) that there exists vector $\nu \in \mathbb{R}^d$ such that

$$\begin{bmatrix} \lambda_q^{(1)}(\tau) \\ \lambda_q^{(1)}(\tau) \\ \lambda_q^{(2)}(\tau) \\ \lambda_q^{(2)}(\tau) \end{bmatrix} = \begin{bmatrix} I_d \\ 0_d \\ 0_d \\ 0_d \end{bmatrix} \nu = \begin{bmatrix} \nu \\ 0_d \\ -\nu \\ 0_d \end{bmatrix}. \quad (104)$$

Therefore, it is concluded that $\nu = \lambda_p^{(1)}(\tau) = -\lambda_p^{(2)}(\tau)$ and $\lambda_q^{(1)}(\tau) = \lambda_q^{(2)}(\tau) = 0_d$. This shows that $\lambda_p^{(1)}(t) = (\tau - t)\lambda_p^{(2)}$ and $\lambda_q^{(2)}(t) = (\tau - t)\lambda_q^{(1)}$. Note that if one has $\lambda_p^{(1)} = 0_d$ or $\lambda_p^{(2)} = 0_d$, then $\nu = 0_d$ and $\lambda_q^{(1)} = \lambda_q^{(2)} = 0_d$. Also, from (97) and (101), it is concluded that $H = \lambda_0$ and $\lambda_0 = 0$. Therefore, for any $t \in [0, \tau]$, one has $(\lambda_0, \lambda^{(1)}(t), \lambda^{(2)}(t)) = 0$, which is not possible due to the initial assumptions on $\lambda_0$, $\lambda^{(1)}$ and $\lambda^{(2)}$. This shows that $\lambda_p^{(1)} \neq 0_d$ and $\lambda_p^{(2)} \neq 0_d$. Considering $\lambda_q^{(1)}(t) = (\tau - t)\lambda_p^{(1)}$ and $\lambda_q^{(2)}(t) = (\tau - t)\lambda_p^{(2)}$, and also the equation (99), it can be verified that optimal control inputs $u^{(1)}$ and $u^{(2)}$ are constant functions such that

$$u^{(1)}(t) = -\frac{\nu}{\|\nu\|} m_2, \quad u^{(2)}(t) = \frac{\nu}{\|\nu\|} (m_2 - e), \quad (105)$$

for any $t \in [0, \tau]$. Therefore, one has

$$2m_2 - \epsilon \frac{\nu}{\|\nu\|} \tau^2 + (q - v)\tau + p = 0, \quad (106)$$

and subsequently,

$$\frac{1}{2}(2m_2 - \epsilon)\tau^2 = \|(q - v)\tau + p - t\|. \quad (107)$$

Therefore, $\hat{t} \leq \tau$. This says that $\hat{t}$ is smaller than the earliest time that a vehicle starting from point $p$ with velocity $q$ can visit a target which is now at the point $r$ and moving with velocity $p$. Therefore, the vehicle does not visit the target prior to time $\hat{t}$.

For any $t \in [0, \hat{t}]$, let $f_t$ be a function as $f_t : \mathbb{R}_{\geq 0} \to \mathbb{R}^d$ with

$$f_t(s) = \frac{1}{2}\dot{\gamma}(t)s^2 + (\gamma(t) - ut - q)s + \gamma(t) - \frac{1}{2}ut^2 + pt + p \| \frac{m_2}{2} \|^2. \quad (108)$$

From triangle inequality, one has

$$f_t(s) \leq \frac{1}{2}\|\dot{\gamma}(t)\|^2 s^2 + \|\dot{\gamma}(t) - ut - q\|s + \gamma(t) - \frac{1}{2}ut^2 + pt + p \| \frac{m_2}{2} \|^2, \quad (109)$$

and subsequently,

$$f_t(s) \leq -\frac{1}{2}\|\dot{\gamma}(t)\|^2 s^2 + \|\gamma(t) - ut - q\|s + \gamma(t) - \frac{1}{2}ut^2 + pt + p \| \frac{m_2}{2} \|^2. \quad (110)$$

From this, one can conclude that $\lim_{s \to \infty} f_t(s) = -\infty$ and consequently, there exists $\hat{s} \in \mathbb{R}_{\geq 0}$ such that $f_t(\hat{s}) < 0$. Also, one has $f_t(0) = \|\gamma(t) - \frac{1}{2}ut^2 - pt + p\| \geq 0$. Since $f_t$ is sum of a polynomial and norm of a vector-value polynomial function which are continuous functions, it is deduced that $f_t$ is a continuous function. Therefore, there exists $\bar{s} \in \mathbb{R}_{\geq 0}$ such that $f_t(\bar{s}) = 0$. Accordingly, set $u(t)$ and $\tau(t)$ as

$$\begin{cases} u(t) = \frac{2}{\tau(t)}[\frac{1}{2}\gamma(t)s^2 + (\gamma(t) - ut - q)s + \gamma(t) - \frac{1}{2}ut^2 + pt + p], \\ \tau(t) = \bar{s}. \end{cases} \quad (111)$$

Therefore, $(\tau(t), u(t))$ is solution for system of equations (49) in $\mathbb{R}_{\geq 0} \times \mathbb{R}^d$. Similarly, one can show that there exist a solution like $(\hat{t}(t), \hat{u}(t))$ in $\mathbb{R}_{\geq 0} \times \mathbb{R}^d$ for system of equations introduced in (50).

For any $t \in [0, \hat{t}]$, one has

$$\begin{cases} \gamma(t) = r + vt + \frac{1}{2}at^2, \\ \gamma(t) = v + at, \end{cases} \quad (112)$$

Therefore, one has

$$\begin{cases} \frac{1}{2}\gamma(t)(\hat{t}(0) - t)^2 = \frac{1}{2}a(t)(0)^2 - a(t)t + \frac{1}{2}at^2, \\ \{\gamma(t) - \hat{u}(0)\} - \gamma(t) - q(\hat{t}(0) - t) = (v + at - \hat{u}(0)(t) - q(\hat{t}(0) - t) = (v - q)t + \hat{u}(0)t^2 + at(0) - \frac{1}{2}
\end{cases}$$

Obtaining the sum of right-hand-sides and left-hand-sides of above equations, one can see that

$$\frac{1}{2}\gamma(t)(\hat{t}(0) - t)^2 + (\gamma(t) - \hat{u}(0) - q)(\hat{t}(0) - t) + (\gamma(t) - \frac{1}{2}\hat{u}(0)t^2 - q\hat{t}(0) - p) = \frac{1}{2}\hat{u}(0)(\hat{t}(0) - t)^2. \quad (113)$$

Since $t \leq \hat{t}$ and $\hat{t} \leq \hat{t}(0)$, one has $t \leq \hat{t}(0)$. Hence, from $\|\hat{u}(0)\| = m_2$ and the definition of $(\hat{t}(0), \hat{u}(0))$ and $(\hat{t}(0), \hat{u}(0))$, it is concluded that

$$(\hat{t}(0), \hat{u}(0)) = (\hat{t}(0) - t, \hat{u}(0)). \quad (114)$$

Define functions $\delta_t, \delta_t, \delta_n, \delta_n \in \mathbb{R}^d$ and $\delta_r : [0, \hat{t}] \to \mathbb{R}$ as following

$$\begin{cases} \delta_t(\hat{r}) = \gamma(\hat{r}) - \gamma(t), \\ \delta_r(\hat{r}) = \gamma(\hat{r}) - \gamma(t), \\ \delta_n(\hat{r}) = \gamma(\hat{r}) - \gamma(t), \\ \delta_n(\hat{r}) = \gamma(\hat{r}) - \gamma(t), \end{cases} \quad (115)$$

Note that

$$\begin{cases} \delta_t(\hat{r}) = \frac{1}{2}a(t)^2, \\ \delta_r(\hat{r}) = \frac{1}{2}a^2(t)^2 + \frac{1}{2}a(t)^2, \\ \delta_n(\hat{r}) = \frac{1}{2}a^2(t)^2 + \frac{1}{2}a(t)^2 + \frac{1}{2}a(t)^2, \end{cases} \quad (116)$$

Accordingly, from (49) it is concluded that

$$\begin{cases} \frac{1}{2}\hat{u}(0) + \delta_n(\hat{r}) = \frac{1}{2}\gamma(\hat{r}) + \delta_n(\hat{r}) = \frac{1}{2}\gamma + \delta_n(\hat{r}) + \delta_n(\hat{r}) + \gamma(\hat{r}) + \delta_n(\hat{r}) + \frac{1}{2}a(t)^2, \\ \frac{1}{2}\hat{u}(0) + \delta_n(\hat{r}) = \frac{1}{2}m_2. \end{cases} \quad (117)$$
where \( \tilde{p} = \frac{1}{2} \dot{u}(0) h^2 + q h + p \) and \( \tilde{q} = \frac{1}{2} \dot{u}(0) + q. \) Also, recall that
\[
\frac{1}{2} \frac{n_u}{\|u\|^2} = \frac{1}{2} n_u^2.
\]

Now, assume that \( \tilde{u}^T (\tilde{a} \tilde{\tau}(0) + v - \tilde{u}(0) \tilde{\tau}(0) - q) \neq 0. \) Denote by \( \tilde{a}, \tilde{v}, \tilde{r}, \tilde{u}, \) and \( \tilde{\tau} \) as \( a, h + v, \frac{1}{2} h^2 + v h + r, \tilde{u}(0) \) and \( \tilde{\tau}(0) - h, \) respectively. For the given vectors \( p, q \in \mathbb{R}^n, \) define function \( f : \mathbb{R}^{n+n+n} \times \mathbb{R}^n+1 \rightarrow \mathbb{R}^n+1 \) as followings
\[
f(a,v,r,u,\tau) = \left( \frac{1}{2} (u-a)^T \tau^2 + (q-v) \tau + (p-r), \right).
\]
Let \( x \) and \( y \) represent the vectors \( (a^T, v^T, r^T, u^T, \tau^T), \) respectively. Similarly, one can define \( \hat{x} \) and \( \hat{y}. \) One can verify that
\[
D_y f(\hat{x}, \hat{y}) = \left[ I_{n^2} \frac{\tilde{u}^T}{u^T} (\tilde{u} - \tilde{a}) \tilde{\tau} + (\tilde{q} - \tilde{v}) \right],
\]
and
\[
D_x f(\hat{x}, \hat{y}) = \left[ \frac{\tilde{u} \tilde{u}^T}{\tilde{u}^T}, 0_n, 0_n, 0_n, 0_n \right].
\]
Since \( \tilde{u}^T [((\tilde{u} - \tilde{a}) \tilde{\tau} + (\tilde{q} - \tilde{v})) \neq 0, \) it yields \( \det D_y f(\hat{x}, \hat{y}) \neq 0. \) Therefore, using implicit function theorem, it can be seen that there exist a neighborhood of \( \hat{x} \) denoted by \( \mathcal{W} \subseteq \mathbb{R}^{n+n+n} \) and a function \( F : \mathcal{W} \rightarrow \mathbb{R}^n+1 \) such that \( \hat{y} = F(\hat{x}), \) for any \( \hat{x} \in \mathcal{W}. \) Moreover, one has
\[
\hat{y} = F(\hat{x}) + D_x F(\hat{x})(\hat{x} - \hat{x}) + O(\|\hat{x} - \hat{x}\|^2)
\]
where \( f(\hat{x}, \hat{y}) = 0 \) and \( D_x F(\hat{x}) = -(D_y f(\hat{x}, \hat{y}))^{-1} D_x f(\hat{x}, \hat{y}). \)

Regarding \( D_x f(\hat{x}), \) one has
\[
D_x F(\hat{x}) = \left[ \frac{\tilde{u} \tilde{u}^T}{\tilde{u}^T} \right]
\]
where \( \tilde{w} \) denotes \( \frac{1}{2} ((\tilde{u} - \tilde{a}) \tilde{\tau} + (\tilde{q} - \tilde{v})); \) and the matrix inversion in (119) is obtained as following:
\[
\frac{1}{2} \frac{\tilde{u} \tilde{u}^T}{\tilde{u}^T} = \left[ \frac{\tilde{w} \tilde{w}^T}{\tilde{w}^T} \right]^{-1} = \frac{1}{\tilde{w}^T \tilde{w}} \tilde{w} \tilde{w}^T.
\]
According to the above discussion, one has
\[
\begin{bmatrix}
\delta_u \\
\delta_{\tau}
\end{bmatrix} = \begin{bmatrix}
I_n - \frac{\tilde{u} \tilde{u}^T}{\tilde{u}^T} \\
\frac{1}{\tilde{u}^T} \tilde{u}^T
\end{bmatrix} \left[ \frac{\tilde{u} \tilde{u}^T}{\tilde{u}^T} \right] \left[ \frac{\tilde{w} \tilde{w}^T}{\tilde{w}^T} \right] \begin{bmatrix}
\delta_u \\
\delta_{\tau}
\end{bmatrix} + O(h^2),
\]
From (112), it is concluded that
\[
\begin{bmatrix}
\delta_u (h) \\
\delta_{\tau} (h)
\end{bmatrix} = \left[ I_n - \frac{\tilde{u} \tilde{u}^T}{\tilde{u}^T} \right] \begin{bmatrix}
\hat{u} (h) \\
\hat{\tau} (h)
\end{bmatrix} + o(h^2),
\]
and subsequently,
\[
\begin{bmatrix}
\frac{\delta_u (h)}{h} \\
\frac{\delta_{\tau} (h)}{h}
\end{bmatrix} = \left[ I_n - \frac{\tilde{u} \tilde{u}^T}{\tilde{u}^T} \right] \begin{bmatrix}
\hat{u} (h) \\
\hat{\tau} (h)
\end{bmatrix} + o(h).
\]