Image-based Performance Certificate for Autonomous Ground Vehicles

Haoming Jing[1†], Maitham AL-Sunni[1†], Yorie Nakahira[1*]

***Abstract*— This paper focuses on assuring the performance of autonomous grand vehicles with monocular vision-based controllers. Significant progress has been made for perception-based control techniques when accurate state estimation, depth information, and stereo vision or LiDAR are available. However, achieving provable guarantees with monocular vision-based controllers remains a challenging task. In this study, our goal is to create monocular image-based performance certificates that do not depend on precise state and depth estimates. The performance specifications take the form of forward invariance or forward convergence conditions of the system state in the vehicle coordinate. To achieve this, we first translate the performance specifications into monocular image sequence conditions and then determine the set of admissible control actions that meet these conditions. Characterizing sufficient control conditions using system dynamics in vehicle coordinates often results in the set of admissible control actions being parameterized by system state and distance information. In contrast, the proposed approach allows the set of admissible control actions to be directly computed using monocular image information. The effectiveness of the proposed method is demonstrated using simulation.**

I. Introduction

Autonomous systems often require operating within strict size, weight, and power constraints, which may lead to the use of learning-based techniques that rely on image sequences captured by cameras attached to the vehicle [1]. However, some autonomous systems are equipped with only a monocular camera, which limits their depth perception. This, coupled with the black-box nature of image-based controllers and the absence of accurate depth and range sensors, creates significant challenges in ensuring safety and achieving desired targets [2], [3], [4], [5]. Therefore, addressing these challenges is essential to assure the safety and reliability of autonomous systems.

This paper focuses on providing provable safety and performance in perception-based controllers. We consider the settings when the control of an autonomous ground vehicle is based on a single camera and easily measurable ego vehicle state (e.g., speed and turning rate using a speedometer and the steering specifications for the specific vehicle). We do not assume the availability of state and distance estimates in the global/vehicle coordinate systems or their errors are bounded. We investigate the conditions on the image sequences that ensure the forward invariance and forward convergence conditions of the state in vehicle coordinates. We then characterize the set of admission control actions that ensures these image sequence conditions. The proposed methods can be used to certify or guarantee behaviors such as target tracking and obstacle avoidance, which are also demonstrated using numerical simulation. These proposed methods can be modularly added to a nominal controller (for example, an end-to-end neural network controller) without requiring explicit construction of state or distance estimate.

A. Related work

1) End-to-end black-box controllers: Many end-to-end learning techniques are developed for autonomous systems that use cameras and other sensor information to compute actions without explicitly constructing the perception, planning, and control modules. This topic has been extensively studied for diverse applications such as autonomous ground vehicles, unmanned aerial vehicles, and robotics. Common approaches include deep reinforcement learning methods [6], [7], [8], supervised deep learning methods [9], [10], and imitation learning [11]. These approaches often give black-box control which may not come with guaranteed behaviors. Our characterization of the admissible set of control actions can be used to certify or modify the output of such controllers for assured safety and performance.

2) Visual servoing: Many techniques are developed to control the motion of robots or unmanned aerial vehicles using visual feedback. These techniques consider diverse settings such as the use of stereo or monocular cameras, LiDARs, and depth sensors. These techniques can take a position-based approach and an image-based approach [12]. The former uses image features to estimate the position and states of objects which is then used to minimize the error between the current and desired state. The latter defines certain image features and minimizes the errors between the current and desired features. These techniques are used to perform a variety of tasks such as visual tracking and obstacle avoidance [13], [14]. While these techniques aim at designing the most desirable control action (or policy), the proposed methods characterize the set of admissible actions with guaranteed performance, which can be used to modify any nominal (or black-box) controller.

3) Human driving behaviors: There is extensive research that studies the behavior of human drivers. Human drivers are known to decide their control action based on the information of visual information without estimating the state in the global or vehicle coordinate system. [15], [16], [17]). Experiments on such control strategies have been done using optical flow-based focus expansion with virtual data [18] and by using cameras [15], [16], [17]. Many models have been developed to characterize such policies [19], [20]. This

†These authors contributed equally.

[1]The authors are with the Department of Electrical and Computer Engineering, Carnegie Mellon Universty, {haomingj,malsunni, ynakahir}@andrew.cmu.edu.

∗To whom correspondence should be addressed.

work is inspired by these studies, with additional provable performance and safety guarantees.

4) Robust and safe control techniques: Significant progress has been made for perception-based control techniques when state estimation with bounded errors, depth information, or stereo vision or Lidar are available. Control barrier functions to constrain the output of perception-based controllers [21], [22], [23], [24]. Perception-based control problem has also been formulated into a linear system with bounded state estimation errors, on which system-level synthesis can be used to guarantee robust performance [25], [26]. While these techniques primarily focus on assuring safety given the observation processes and state estimates, this paper investigates the types of specifications that can be ensured using a monocular camera as well as the image features that are sufficient to characterize the admissible set of control actions.

II. Problem Statement

In this paper, we consider an autonomous ground vehicle system. The system uses image input for perception and calculates the appropriate control actions that drive the vehicle.

A. System Description

We consider the class of autonomous ground vehicle systems whose dynamics can be generalized using the following dynamical system:

$$\begin{aligned} &\dot{x}_t^e = v_t^e \cos(\theta_t^e),\ \dot{y}_t^e = v_t^e \sin(\theta_t^e),\ \dot{z}_t^e = 0, \\ &\dot{v}_t^e = a_t^e,\ \dot{\theta}_t^e = \psi_t^e,\ \dot{\psi}_t^e = \phi_t^e, \end{aligned} \tag{1}$$

where $[x^e\ y^e\ z^e\ v^e\ \theta^e\ \psi^e]^T := x \in \mathbb{R}^6$ is the state of the vehicle and $[a^e\ \phi^e]^T := u \in \mathbb{R}^2$ is the control action. Here, the state x contains the global coordinate information (*e.g.,* locations and velocities) of the ego vehicle. The location of the vehicle is determined by x^e and y^e, and z^e, the height of the ego vehicle, is assumed to be constant. There are 2 control actions: the acceleration a and the changing rate of turning rate ϕ. Let $\mathcal{U}_s$ be the set of feasible control actions. The control actions are determined based on the sensor information. We assume a nominal control action u_t^{nom} is available at all times, which achieves some nominal performance of the system without considering the control objective. The vehicle has cameras as sensors, which provide sequential images without depth information. Let the camera output at each frame a $m_c \times n_c$ image. Let

$$o_t^i := \begin{bmatrix} X_t^i \\ Y_t^i \end{bmatrix} \tag{2}$$

be the location of pixel i on the image at time t. Here, we use the image coordinate system introduced in [16], whose coordinate transformation is given by

$$\begin{bmatrix} X_t^i \\ Y_t^i \end{bmatrix} = f \begin{bmatrix} \frac{y_t^{\mathrm{v}i}}{x_t^{\mathrm{v}i}} \\ \frac{z_t^{\mathrm{v}i}}{x_t^{\mathrm{v}i}} \end{bmatrix}. \tag{3}$$

Here, $x_t^{\mathrm{v}i}$, $y_t^{\mathrm{v}i}$ and $z_t^{\mathrm{v}i}$ are the coordinates of the pixel i in the vehicle coordinate at time t[1]. The camera is located at $x^{\mathrm{v}}, y^{\mathrm{v}}, z^{\mathrm{v}} = 0$, and the focal length of the camera is f. In this paper, we only consider objects located in front of the ego vehicle and below the camera height, *i.e.,* $x^{\mathrm{v}} > 0$ and $z^{\mathrm{v}} > 0$ for all objects, where the heights of the objects do not change.

An optical flow containing the information about where pixels move is defined as the time derivative of the image coordinate locations[2]. Let

$$q_t^i := \begin{bmatrix} \mu_t^i \\ \nu_t^i \end{bmatrix} := \begin{bmatrix} \dot{X}_t^i \\ \dot{Y}_t^i \end{bmatrix} \tag{4}$$

denote the optical flow of the pixel i. Let $\mathcal{O}_t^i := \{X_t^i, Y_t^i, \mu_t^i, \nu_t^i\}$. Let $\mathcal{O}_t := \bigcup_{i\in\{1,2,\cdots,m_c n_c\}} \mathcal{O}_t^i$. Let $\mathcal{Q}_t^i := \mathcal{O}_t^i \cup \mathcal{V}_t$ and $\mathcal{Q}_t := \mathcal{O}_t \cup \mathcal{V}_t$, where $\mathcal{V}_t$ includes easily measurable state information without requiring for global coordinate information (*e.g.,* ego vehicle speed and turning rate) at time t. The information available to the controller at time t is $\{\mathcal{Q}_\tau : \tau \le t\}$.

B. Design Objectives

The design objective is to achieve common control objectives in ground vehicle/robot systems using information $\mathcal{Q}$ only, while minimizing deviation from the nominal control action u^{nom}. The desired control action is given by

$$\begin{aligned} u_t = \quad & \arg\min_u \quad J(u_t^{nom}, u) \\ & \text{s.t.} \quad u \text{ satisfying control objectives,} \end{aligned} \tag{5}$$

where $J : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ is a function that penalizes derivation from u^{nom}. Conventionally, such control objectives, including object/trajectory tracking and obstacle avoidance, require global information to ensure the objectives. Specifically, the objectives are characterized in the global coordinate system using the global coordinate location of the ego vehicle and the objects in the surrounding environment. To find these locations, the system must perform localization based on observations with depth information, which is subject to large observation errors in adversarial conditions. To avoid this critical dependence on depth information and use more reliable information to perform control, we would like to find conditions that can be certified with the control action and information $\mathcal{Q}$, *i.e.,* image coordinates information, and easily measurable ego vehicle states. Then, we would like to translate the conditions characterized using global coordinate states to these conditions, so that conditions characterized using global coordinate states can also be certified with the control action and information $\mathcal{Q}$. Such certification

[1]The vehicle coordinate system is a 3-dimensional Cartesian coordinate system where the ego vehicle is at the origin and traveling to the positive x-axis direction. The z-axis points to the downward direction.

[2]Here, we make a slight misuse of the notion 'pixel', since the pixel will not move on the image. Instead, what moves is the content within the pixel. For the rest of the paper, we will use the word 'pixel' to refer to the content within that pixel. To simplify the notation, we also make the content in one pixel to not lose its index i when we perform control using information from different times, although the content has moved to a different pixel during the time interval.

process can be added to existing methods such as end-to-end autonomous driving algorithms to provide additional performance guarantees.

III. Proposed Method

In this section, we first introduce sufficient control conditions to ensure some properties in the optical flow. Then, using a few application examples, we introduce the relationship between the conditions in vehicle coordinates and the conditions in the image plane, such that some conditions in the vehicle coordinates can be satisfied by guaranteeing certain conditions in the image plane. Finally, we propose an algorithm that practically finds the desired control action using (5).

A. Optical Flow Properties Guarantee

We will find the sufficient conditions in the control action to guarantee some properties in the optical flow without the use of any global coordinate information. Below, we show that there exist conditions that we can use to ensure that μ_t^i and ν_t^i are within unbounded intervals, *i.e.,*

$$\mu^i > \mu_b^i, \tag{6}$$
$$\mu^i < \mu_b^i, \tag{7}$$
$$\nu^i > \nu_b^i, \text{ or} \tag{8}$$
$$\nu^i < \nu_b^i. \tag{9}$$

We will show later using examples in Section III-B that conditions (6) to (9) are very useful in many common autonomous driving tasks, including target following and obstacle avoidance.

Below, we will present some theorems that show (6) to (9) can be ensured using only $\mathcal{Q}$, *i.e.,* information available in the image coordinates, under the reasonable assumptions that, the speed and turning rate of all pixels relevant to the control objective are bounded. We use $\mathcal{I}_t$ to denote the set of pixels relevant to the control objective at time t. We identify such pixels using common perception algorithms, from which we don't require depth information. The assumptions can be written as

$$\underline{v}^i \le v_t^i \le \overline{v}^i, \underline{\psi}^i \le \psi_t^i \le \overline{\psi}^i, \underline{a}^i \le a_t^i \le \overline{a}_t^i, \ \forall t \ge 0, i \in \mathcal{I}_t. \tag{10}$$

Here, v^i, ψ^i and a^i denote the speed and turning rate of pixel i, respectively, and $\underline{v}^i$, $\underline{\psi}^i$, $\overline{v}^i$, $\overline{\psi}^i$, $\underline{a}^i$ and $\overline{a}^i$ denote their lower bounds and upper bounds, respectively. Before introducing the theorems, we first define some equations and Algorithm 1. We will show the derivation for these equations in the proofs for Theorems 1 and 2. Let

$$\begin{aligned}\dot{\mu}_t^i = {} & \phi_t^e f + \frac{2\mu_t^i \nu_t^i}{Y_t^i} + \frac{(X_t^i)^2 \phi_t^e}{f} + \frac{X_t^i a_t^e}{D_t^i \mathcal{X}_t^i} + \frac{a_t^i f \sin(\Theta_t^i)}{D_t^i \mathcal{X}_t^i} \\ & - \frac{f\psi_t^e v_t^e}{D_t^i \mathcal{X}_t^i} - \frac{X_t^i a_t^i \cos(\Theta_t^i)}{D_t^i \mathcal{X}_t^i} + \frac{2f\psi_t^e v_t^i \cos(\Theta_t^i)}{D_t^i \mathcal{X}_t^i} \\ & - \frac{f\psi_t^i v_t^i \cos(\Theta_t^i)}{D_t^i \mathcal{X}_t^i} + \sin(\Theta_t^i)\left(\frac{2X_t^i \psi_t^e v_t^i}{D_t^i \mathcal{X}_t^i} - \frac{X_t^i \psi_t^i v_t^i}{D_t^i \mathcal{X}_t^i}\right),\end{aligned} \tag{11}$$

where $\Theta^i = (\theta^e - \theta^i)$ is the difference between the traveling direction of the ego vehicle and pixel i in global coordinates, D^i is the distance between the content in pixel i and the ego vehicle, and

$$\mathcal{X}_t^i = \frac{1}{\sqrt{\frac{(X_t^i)^2}{f^2} + 1}}. \tag{12}$$

Let

$$\begin{aligned}\dot{\nu}_t^i = {} & Y_t^i (\psi_t^e)^2 + \frac{2Y_t^i v_t^i \sin(\Theta_t^i)\psi_t^e}{D_t^i \mathcal{X}_t^i} + \frac{2(\nu_t^i)^2}{Y_t^i} + \frac{Y_t^i a_t^e}{D_t^i \mathcal{X}_t^i} \\ & + \frac{(X_t^i Y_t^i \phi_t^e)}{f} - \frac{(Y_t^i a_t^i \cos(\Theta_t^i))}{D_t^i \mathcal{X}_t^i} - \frac{Y_t^i \psi_t^i v_t^i \sin(\Theta_t^i)}{D_t^i \mathcal{X}_t^i},\end{aligned} \tag{13}$$

$$D_t^i(\mu_t^i) = -\frac{f(X_t^i v_t^e - X_t^i v_t^i \cos(\Theta_t^i) + f v_t^i \sin(\Theta_t^i))}{\mathcal{X}_t^i(\psi_t^e (X_t^i)^2 + \psi_t^e f^2 - \mu_t^i f)}, \tag{14}$$

and

$$D_t^i(\nu_t^i) = \frac{Y_t^i f(v_t^e - v_t^i \cos(\Theta_t^i))}{\mathcal{X}_t^i(f\nu_t^i - X_t^i Y_t^i \psi_t^e)}. \tag{15}$$

In Algorithm 1, we first compute[3] a bound for D^i using μ^i and ν^i. Then, we use the bound to compute the bound for $\dot{\mu}^i$ and $\dot{\nu}^i$. Using these bounds, we verify if the given control action u satisfies the required conditions for control objectives.

Algorithm 1 Control action verification

Require: Desired boundary value for optical flow μ_b^i
Require: Information $\mathcal{Q}_t^i = \{X_t^i, Y_t^i, \mu_t^i, \nu_t^i, v_t^e, \psi_t^e\}$
Require: Lower bounds $\underline{v}^i$, $\underline{\psi}^i$ and upper bounds $\overline{v}^i$, $\overline{\psi}^i$ of the speed and turning rate of the pixel i at time t

procedure VerifyControlAction(u_t,i)
  Calculate expressions (11), (13), (14), and (15), to get $\dot{\mu}_t^i$, $\dot{\nu}_t^i$, $D_t^i(\mu_t^i)$, and $D_t^i(\nu_t^i)$,[4] respectively, using $\mathcal{Q}_t^i$ and u_t
  $\underline{D}_t^{\mu i} \leftarrow \min_{\underline{v}^i \le v_t^i \le \overline{v}^i, 0 \le \Theta_t < 2\pi} D_t^i(\mu_t^i)$
  $\overline{D}_t^{\mu i} \leftarrow \max_{\underline{v}^i \le v_t^i \le \overline{v}^i, 0 \le \Theta_t < 2\pi} D_t^i(\mu_t^i)$
  $\underline{D}_t^{\nu i} \leftarrow \min_{\underline{v}^i \le v_t^i \le \overline{v}^i, 0 \le \Theta_t < 2\pi} D_t^i(\nu_t^i)$
  $\overline{D}_t^{\nu i} \leftarrow \max_{\underline{v}^i \le v_t^i \le \overline{v}^i, 0 \le \Theta_t < 2\pi} D_t^i(\nu_t^i)$
  $\underline{D}_t^i \leftarrow \max\{\underline{D}_t^{\mu i}, \underline{D}_t^{\nu i}\}$
  $\overline{D}_t^i \leftarrow \min\{\overline{D}_t^{\mu i}, \overline{D}_t^{\nu i}\}$
  if $\mu^i < \mu_b^i$ is desired and $\mu_t^i \ge \mu_b^i$ **then**
    $\overline{\dot{\mu}}_t^i \leftarrow \max_{\substack{\underline{v}^i \le v_t^i \le \overline{v}^i, 0 \le \Theta_t < 2\pi, \underline{D}_t^i \le D_t^i \le \overline{D}_t^i, \\ \underline{\psi}^i \le \psi_t^i \le \overline{\psi}^i, \underline{a}^i \le a_t^i \le \overline{a}^i}} \dot{\mu}_t^i$
    if $\overline{\dot{\mu}}_t^i \ge 0$ **then**
      u_t is not certified for μ
    else

[3] In some obstacle avoidance settings, a lower bound D^i is readily available if it is desired that the obstacle be kept outside a certain radius of the ego vehicle, and such objective is ensured using a forward invariance condition.

[4] To distinguish between (14) and (15), we use $D_t^i(\mu_t^i)$ to denote the expression for D_t^i derived from (14) with μ_t^i, and $D_t^i(\nu_t^i)$ to denote the expression for D_t^i derived from (15) with ν_t^i.

u_t is certified for μ
end if
else if $\mu^i > \mu_b^i$ is desired and $\mu_t^i \leq \mu_b^i$ **then**
$\underline{\dot{\mu}}_t^i \leftarrow \min_{\underline{v}^i \leq v_t^i \leq \overline{v}^i, 0 \leq \Theta_t < 2\pi, \underline{D}_t^i \leq D_t^i \leq \overline{D}_t^i, \underline{\psi}^i \leq \psi_t^i \leq \overline{\psi}^i, \underline{a}^i \leq a_t^i \leq \overline{a}^i} \dot{\mu}_t^i$
if $\underline{\dot{\mu}}_t^i \leq 0$ **then**
u_t is not certified for μ
else
u_t is certified for μ
end if
else
u_t is certified for μ
end if
if $\nu^i < \nu_b^i$ is desired and $\nu_t^i \geq \nu_b^i$ **then**
$\overline{\dot{\nu}}_t^i \leftarrow \max_{\underline{v}^i \leq v_t^i \leq \overline{v}^i, 0 \leq \Theta_t < 2\pi, \underline{D}_t^i \leq D_t^i \leq \overline{D}_t^i, \underline{\psi}^i \leq \psi_t^i \leq \overline{\psi}^i, \underline{a}^i \leq a_t^i \leq \overline{a}^i} \dot{\nu}_t^i$
if $\overline{\dot{\nu}}_t^i \geq 0$ **then**
u_t is not certified for ν
else
u_t is certified for ν
end if
else if $\nu^i > \nu_b^i$ is desired and $\nu_t^i \leq \nu_b^i$ **then**
$\underline{\dot{\nu}}_t^i \leftarrow \min_{\underline{v}^i \leq v_t^i \leq \overline{v}^i, 0 \leq \Theta_t < 2\pi, \underline{D}_t^i \leq D_t^i \leq \overline{D}_t^i, \underline{\psi}^i \leq \psi_t^i \leq \overline{\psi}^i, \underline{a}^i \leq a_t^i \leq \overline{a}^i} \dot{\nu}_t^i$
if $\underline{\dot{\nu}}_t^i \leq 0$ **then**
u_t is not certified for ν
else
u_t is certified for ν
end if
else
u_t is certified for ν
end if
if u_t is certified for μ and ν **then**
return u_t is certified
else
return u_t is not certified
end if
end procedure

Note that Algorithm 1 only uses information $\mathcal{Q}$. Since $\dot{\mu}_t^i$, $\dot{\nu}_t^i$, $D_t^i(\mu_t^i)$, and $D_t^i(\nu_t^i)$ are all nonconvex functions of the optimization variables, we present a lemma about the solution of the optimization problems in Algorithm 1:

Lemma 1. *The global optimal solution can be obtained for the optimization problems in Algorithm 1.*

Given that, we are finally ready to propose the theorems that give a guarantee on conditions (6) to (9):

Theorem 1 (Certification for conditions in μ)**.** *If* (6) *or* (7) *is satisfied at time* 0*, and that* u_t *is certified by Algorithm 1 at all times, then* (6) *or* (7) *is satisfied at all time.*

Theorem 2 (Certification for conditions in ν)**.** *If* (8) *or* (9) *is satisfied at time* 0*, and that* u_t *is certified by Algorithm 1 at all times, then* (8) *or* (9) *is satisfied at all time.*

In the next subsection, we will show that we can ensure some conditions in vehicle coordinates by ensuring certain conditions in the image plane, where the conditions in the image plane can be ensured with the control actions, as shown in the theorems and corollaries above.

B. Ensuring Conditions in Vehicle Coordinates in Image Plane

In this subsection, we will use 2 common tasks in autonomous agent control to demonstrate that, without the use of the vehicle/global coordinate information, certain control specifications in the vehicle coordinate can be ensured by conditions in the image plane. Specifically, in some cases, the condition in vehicle coordinate is equivalent to the condition in the image plane, even though the gaze coordinates transformation (3) is not invertible, *i.e.,* the vehicle coordinate information $\{x^{\mathrm{v}}, y^{\mathrm{v}}, z^{\mathrm{v}}\}$ cannot be fully retrieved from information $\mathcal{Q}$. Specifically, we introduce a target following example, which represents a common forward convergence specification, and an obstacle avoidance example, which represents a common forward invariance specification.

1) Target Following: We propose the following theorem that shows the sufficient condition to ensure the target following:

Theorem 3 (Target Following)**.** *Consider a target* i *located in the image plane. Given that*

$$X_t^i \mu_t^i < 0 \tag{16}$$

is satisfied at all time $t > 0$*, then, the ego agent is following the target, i.e.,* $y^{\mathrm{v}_i} \to 0$*.*

Note that in this task, the condition in the vehicle coordinate and the condition in the image plane are equivalent. Depending on the sign of X_t^i, condition (16) can be translated to either $\mu_t^i > 0$ or $\mu_t^i < 0$, which can be satisfied by applying Algorithm 1, according to Theorem 1.

2) Obstacle Avoidance: In this task, the condition in the condition in the vehicle coordinate and the condition in the image plane are not equivalent. Therefore, we first find the sufficient condition in the image plane that guarantees the condition in the vehicle coordinate:

Lemma 2. *Consider a set of positions in the vehicle coordinate characterized by* $r^{\mathrm{v}} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *and*

$$\mathcal{C}^{\mathrm{v}} = \{x^{\mathrm{v}}, y^{\mathrm{v}} : r^{\mathrm{v}}(x^{\mathrm{v}}, y^{\mathrm{v}}) \geq 0\}, \tag{17}$$

and a set of positions in the image plane given by

$$\mathcal{C}^o = \{X, Y : \min_{\overline{h} \leq h \leq \underline{h}} r^{\mathrm{v}}(\frac{Xh}{Y}, f\frac{h}{Y}) \geq 0\}, \tag{18}$$

where $\underline{h} > \overline{h} > 0$*. Let* $\{x^{\mathrm{v}}, y^{\mathrm{v}}, z^{\mathrm{v}}\}$ *be the location of an object in the vehicle coordinate and* o *be the location of the object in the image coordinate. If* $\underline{h} \geq z^{\mathrm{v}} \geq \overline{h}$*, we have*

$$o \in \mathcal{C}^o \Rightarrow \{x^{\mathrm{v}}, y^{\mathrm{v}}\} \in \mathcal{C}^{\mathrm{v}}. \tag{19}$$

Remark 1. In many common obstacle avoidance settings like the example below, the minimization term in (18) usually takes its minimum at either $h = \underline{h}$ or $h = \overline{h}$.

Here, we consider the obstacle avoidance setting that all obstacles under height $\overline{h}$[5] be kept outside of a circle with radius r centered at the ego agent, *i.e.,* the safe set for obstacles is given by

$$\mathcal{C}^{\mathrm{v}} = \{x^{\mathrm{v}}, y^{\mathrm{v}} : (x^{\mathrm{v}})^2 + (y^{\mathrm{v}})^2 - r^2 \geq 0\}. \tag{20}$$

We propose the following theorem that gives the sufficient conditions in optical flow such that $\{x^{\mathrm{v}}, y^{\mathrm{v}}\} \in \mathcal{C}^{\mathrm{v}}$ is satisfied:

Theorem 4 (Obstacle avoidance)**.** *Consider and obstacle i located in the vehicle coordinate at $\{x_t^{\mathrm{v}_i}, y_t^{\mathrm{v}_i}, z_t^{\mathrm{v}_i}\}$ and image plane at o_t^i at time t. Assuming $z_t^{\mathrm{v}_i} \geq \overline{h}$ for all time t and $\{x_0^{\mathrm{v}_i}, y_0^{\mathrm{v}_i}\} \in \mathcal{C}^v$, given that*

$$\overline{h}^2 X_t^i \mu_t^i - r^2 Y_t^i \nu_t^i > 0 \tag{21}$$

is satisfied whenever

$$(X_t^i)^2 \overline{h}^2 - (Y_t^i)^2 r^2 + \overline{h}^2 f^2 \leq 0, \tag{22}$$

we have $\{x_t^{\mathrm{v}_i}, y_t^{\mathrm{v}_i}\} \in \mathcal{C}^{\mathrm{v}}$ at all time $t > 0$.

Remark 2. Condition (21) can be translated to one of the following sets of conditions:

$$\nu_t^i < \frac{\overline{h}^2 X_t^i \mu_t^i}{r^2 Y_t^i},\ \mu_t^i > \frac{r^2 Y_t^i \nu_t^i}{\overline{h}^2 X_t^i}, \tag{23}$$

or

$$\nu_t^i < \frac{\overline{h}^2 X_t^i \mu_t^i}{r^2 Y_t^i},\ \mu_t^i < \frac{r^2 Y_t^i \nu_t^i}{\overline{h}^2 X_t^i}. \tag{24}$$

In either case, the set of conditions can be satisfied by applying Algorithm 1, according to Theorems 1 and 2.

In the next subsection, we will introduce an algorithm to practically compute the control action based on (5).

C. Computing Control Actions

Here, we discretize the control action space $\mathcal{U}_s$ to a countable set $\mathcal{U}_d$. Assuming $u \in \mathcal{U}_d$, we propose Algorithm 2 to evaluate (5) at time t using Algorithm 1. We assume that there exist at least one control action $u_t \in \mathcal{U}_d$ such that it is certified for all pixel $i \in \mathcal{I}_t$ under Algorithm 1.

IV. Numerical Simulation

In this section, we show the validity of our framework using numerical simulations[6]. We simulate a scenario where we have two objectives, lane-keeping, and obstacle avoidance.

We consider the dynamics given in (1) for both the ego vehicle and the obstacles. The scenario we consider is for an ego vehicle moving on a straight road that contains some obstacles. The objective of the ego vehicle is to avoid obstacles and to stay within the lane as much as possible. The lane-keeping objective is achieved by setting a target point (pixel) in the image plane that corresponds to the middle of the lane. On the other hand, the obstacle avoidance objective is achieved by keeping the track of the obstacle points (the pixels that correspond to the obstacle). The actuation limits are given by $a^e \in [-a_L, a_L]$, and $\phi^e \in [-\phi_L, \phi_L]$.

Algorithm 2 Control action computation

Require: Control action set $\mathcal{U}_d$, nominal control action u_t^{nom}

```
b ← 0        ▷ binary variable indicating whether control action is found
while b < 1 do
    b ← 1
    u_t ← min_{u_t ∈ U_d} J(u^nom, u)
    Remove u_t from U_d
    for i in I_t do
        if VerifyControlAction(u_t, i) is not certified then
            b ← 0
        end if
    end for
end while
if U_d ≠ ∅ then
    u_t is the desired control action at time t
end if
```

We use a proportional controller u_t^{nom} as the nominal controller. The nominal controller governs the vehicle at a constant speed and steers proportionally to the opposite direction of the location of a target point in the image plane such that the target point is pushed to the center of the image but without any guarantees. For the simulator, we start by initializing the environment in the global coordinate system and transform all the information into the vehicle coordinate system. Then, we transform them into the image plane. Afterward, the transformations are done at every time step Δt so that the optical flow information is computed using the discretized version of (4). For the simulated scenario, we assume that the obstacles are static and the velocity of the ego vehicle is constant. We start the environment with the ego vehicle's camera at the global location $(x_0^e, y_0^e, z_0^e) = (0, 0, 1)$, and randomly initialize the global location of the M obstacles such that each obstacle i is defined to be on the ground. The parameters we use in the simulation is as follows: initial ego vehicle velocity v_0^e = 8 *m/s*, initial ego vehicle orientation θ_0^e = 0 *rad*, initial ego vehicle orientation ψ_0^e = 0 *rad/s*, $\overline{h}$ = 1 *m*, maximum acceleration for the ego vehicle a_L = 0 *m/s*^{2}, ϕ_L = ∞ *rad/s*^{2}, $\overline{v}^i = \underline{v}^i = 0$ *m/s*, $\forall i$, Δt = 0.01 *s*, M = 7, $\underline{D}^i = 1$ *m*, $\forall i$, f = 5 *mm*. Figure 1 shows the results of our framework with the parameters listed above with the specific simulated scenario illustrated.

V. Conclusions

In this paper, we constructed monocular image-based performance certificates for autonomous vehicles that do not depend on precise state and depth estimates. The presented performance specifications take the form of forward invariance and forward convergence conditions of the system state in the global coordinate system. We translated the

[5]Since the z-axis is pointing downward, larger value in $\overline{h}$ indicates lower height.

[6]The code for our implementation is available upon request.

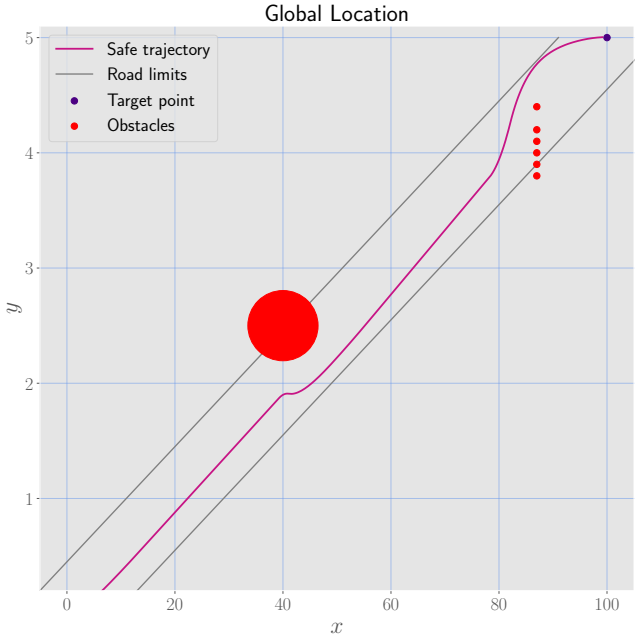

Fig. 1: The trajectory after applying Algorithm 1 with the target following and obstacles avoidance conditions.

performance specifications into monocular image sequence conditions and then determined the set of admissible control actions that meet these conditions. The proposed approach allows the set of admissible control actions to be directly computed using monocular image information.

REFERENCES

[1] A. Tampuu, T. Matiisen, M. Semikin, D. Fishman, and N. Muhammad, "A survey of end-to-end driving: Architectures and training methods," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 4, pp. 1364–1384, 2020.

[2] W. Zeng, W. Luo, S. Suo, A. Sadat, B. Yang, S. Casas, and R. Urtasun, "End-to-end interpretable neural motion planner," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 8660–8669.

[3] Y. Xiao, F. Codevilla, A. Gurram, O. Urfalioglu, and A. M. López, "Multimodal end-to-end autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 1, pp. 537–547, 2020.

[4] M. Bojarski, P. Yeres, A. Choromanska, K. Choromanski, B. Firner, L. Jackel, and U. Muller, "Explaining how a deep neural network trained with end-to-end learning steers a car," *arXiv preprint arXiv:1704.07911*, 2017.

[5] J. Kim and J. Canny, "Interpretable learning for self-driving cars by visualizing causal attention," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2942–2950.

[6] J. Chen, S. E. Li, and M. Tomizuka, "Interpretable end-to-end urban autonomous driving with latent deep reinforcement learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 5068–5078, 2021.

[7] Z. Huang, J. Zhang, R. Tian, and Y. Zhang, "End-to-end autonomous driving decision based on deep reinforcement learning," in *2019 5th International Conference on Control, Automation and Robotics (ICCAR)*, 2019, pp. 658–662.

[8] M. Jaritz, R. De Charette, M. Toromanoff, E. Perot, and F. Nashashibi, "End-to-end race driving with deep reinforcement learning," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 2070–2075.

[9] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang *et al.*, "End to end learning for self-driving cars," *arXiv preprint arXiv:1604.07316*, 2016.

[10] H. Xu, Y. Gao, F. Yu, and T. Darrell, "End-to-end learning of driving models from large-scale video datasets," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2174–2182.

[11] F. Codevilla, M. Müller, A. López, V. Koltun, and A. Dosovitskiy, "End-to-end driving via conditional imitation learning," in *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2018, pp. 4693–4700.

[12] F. Chaumette and S. Hutchinson, "Visual servoing and visual tracking," 2008.

[13] M.-L. Chiang, S.-H. Tsai, C.-M. Huang, and K.-T. Tao, "Adaptive visual servoing for obstacle avoidance of micro unmanned aerial vehicle with optical flow and switched system model," *Processes*, vol. 9, no. 12, p. 2126, 2021.

[14] A. Schaub and D. Burschka, "Reactive avoidance of dynamic obstacles through optimization of their epipoles," in *2015 19th International Conference on System Theory, Control and Computing (ICSTCC)*. IEEE, 2015, pp. 318–324.

[15] Y. Okafuji, T. Fukao, Y. Yokokohji, and H. Inou, "Optical flow-based control for automatic steering systems," in *2015 IEEE/SICE International Symposium on System Integration (SII)*. IEEE, 2015, pp. 483–488.

[16] Y. Okafuji, T. Fukao, and H. Inou, "Development of automatic steering system by modeling human behavior based on optical flow," *Journal of Robotics and Mechatronics*, vol. 27, no. 2, pp. 136–145, 2015.

[17] Y. Okafuji, T. Fukao, Y. Yokokohji, and H. Inou, "Design of a preview driver model based on optical flow," *IEEE Transactions on Intelligent Vehicles*, vol. 1, no. 3, pp. 266–276, 2016.

[18] H. Inou, T. Fukao, S. Totsuka, and Y. Okafuji, "Development of steering control system based on optical flow model," *Transactions of Society of Automotive Engineers of Japan*, vol. 46, no. 2, 2015.

[19] Y. Okafuji, C. D. Mole, N. Merat, T. Fukao, Y. Yokokohji, H. Inou, and R. M. Wilkie, "Steering bends and changing lanes: the impact of optic flow and road edges on two point steering control," *Journal of vision*, vol. 18, no. 9, pp. 14–14, 2018.

[20] Y. Okafuji, T. Fukao, and H. Inou, "Theoretical interpretation of driver's gaze considering optic flow and seat position," *IFAC-PapersOnLine*, vol. 52, no. 19, pp. 335–340, 2019.

[21] D. Kalaria, Q. Lin, and J. M. Dolan, "Towards safety assured end-to-end vision-based control for autonomous racing," *arXiv preprint arXiv:2303.02267*, 2023.

[22] W. Xiao, T.-H. Wang, M. Chahine, A. Amini, R. Hasani, and D. Rus, "Differentiable control barrier functions for vision-based end-to-end autonomous driving," *arXiv preprint arXiv:2203.02401*, 2022.

[23] S. Dean, A. Taylor, R. Cosner, B. Recht, and A. Ames, "Guaranteeing safety of learned perception modules via measurement-robust control barrier functions," in *Conference on Robot Learning*. PMLR, 2021, pp. 654–670.

[24] F. Castañeda, H. Nishimura, R. McAllister, K. Sreenath, and A. Gaidon, "In-distribution barrier functions: Self-supervised policy filters that avoid out-of-distribution states," *arXiv preprint arXiv:2301.12012*, 2023.

[25] S. Dean, N. Matni, B. Recht, and V. Ye, "Robust guarantees for perception-based control," in *Learning for Dynamics and Control*. PMLR, 2020, pp. 350–360.

[26] H. Zhou and V. Tzoumas, "Safe perception-based control with minimal worst-case dynamic regret," *arXiv preprint arXiv:2208.08929*, 2022.

[27] M. Nagumo, "Über die lage der integralkurven gewöhnlicher differentialgleichungen," *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series*, vol. 24, pp. 551–559, 1942.

APPENDIX

A. Proofs

Proof (Lemma 1). We first introduce the transformation from global coordinate to vehicle coordinate:

$$\begin{bmatrix} x_t^{\mathrm{v}_i} \\ y_t^{\mathrm{v}_i} \\ z_t^{\mathrm{v}_i} \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta_t^e & \sin\theta_t^e & 0 & -x_t^e\cos\theta_t^e - y_t^e\sin\theta_t^e \\ \sin\theta_t^e & -\cos\theta_t^e & 0 & -x_t^e\sin\theta_t^e + y_t^e\cos\theta_t^e \\ 0 & 0 & 1 & -z_t^e \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t^{\mathrm{g}_i} \\ y_t^{\mathrm{g}_i} \\ z_t^{\mathrm{g}_i} \\ 1 \end{bmatrix}, \tag{25}$$

where the superscript g_i denotes the global coordinate state information of object i. Since

$$\mu_t^i = f\frac{\dot{y}_t^{\mathrm{v}_i}x_t^{\mathrm{v}_i} - y_t^{\mathrm{v}_i}\dot{x}_t^{\mathrm{v}_i}}{(x_t^{\mathrm{v}_i})^2} \tag{26}$$

and

$$\nu_t^i = f\frac{\dot{z}_t^{\mathrm{v}_i}x_t^{\mathrm{v}_i} - z_t^{\mathrm{v}_i}\dot{x}_t^{\mathrm{v}_i}}{(x_t^{\mathrm{v}_i})^2}, \tag{27}$$

we find

$$\begin{aligned} \dot{x}_t^{\mathrm{v}_i} &= \frac{dx_t^{\mathrm{v}_i}}{dx_t^{\mathrm{g}_i}}\frac{dx_t^{\mathrm{g}_i}}{dt} + \frac{dx_t^{\mathrm{v}_i}}{dy_t^{\mathrm{g}_i}}\frac{dy_t^{\mathrm{g}_i}}{dt} + \frac{dx_t^{\mathrm{v}_i}}{d\theta_t^e}\frac{d\theta_t^e}{dt} + \frac{dx_t^{\mathrm{v}_i}}{dx_t^e}\frac{dx_t^e}{dt} + \frac{dx_t^{\mathrm{v}_i}}{dy_t^e}\frac{dy_t^e}{dt} && (28)\\ &= v_t^i\cos(\theta_t^e - \theta_t^i) - \psi_t^e y_t^{\mathrm{v}_i} - v_t^e && (29)\\ &= v_t^i\cos(\Theta_t^i) - \psi_t^e y_t^{\mathrm{v}_i} - v_t^e && (30) \end{aligned}$$

and

$$\begin{aligned} \dot{y}_t^{\mathrm{v}_i} = \frac{dy_t^{\mathrm{v}_i}}{dt} &= \frac{dy_t^{\mathrm{v}_i}}{dx_t^{\mathrm{g}_i}}\frac{dx_t^{\mathrm{g}_i}}{dt} + \frac{dy_t^{\mathrm{v}_i}}{dy_t^{\mathrm{g}_i}}\frac{dy_t^{\mathrm{g}_i}}{dt} + \frac{dy_t^{\mathrm{v}_i}}{d\theta_t^e}\frac{d\theta_t^e}{dt} + \frac{dy_t^{\mathrm{v}_i}}{dx_t^e}\frac{dx_t^e}{dt} + \frac{dy_t^{\mathrm{v}_i}}{dy_t^e}\frac{dy_t^e}{dt} && (31)\\ &= \psi_t^e x_t^{\mathrm{v}_i} + v_t^i\sin(\theta_t^e - \theta_t^i) && (32)\\ &= \psi_t^e x_t^{\mathrm{v}_i} + v_t^i\sin(\Theta_t^i). && (33) \end{aligned}$$

Given that $(D_t^i)^2 = (x_t^{\mathrm{v}_i})^2 + (y_t^{\mathrm{v}_i})^2$, $X_t^i = f\frac{y_t^{\mathrm{v}_i}}{x_t^{\mathrm{v}_i}}$, and $x_t^{\mathrm{v}_i} < 0$, we have

$$x_t^{\mathrm{v}_i} = \frac{D_t^i}{\sqrt{\frac{(X_t^i)^2}{f^2} + 1}} = D_t^i\mathcal{X}_t^i \tag{34}$$

and

$$y_t^{\mathrm{v}_i} = \frac{X_t^i}{f}D_t^i\mathcal{X}_t^i. \tag{35}$$

Therefore, we have

$$\mu_t^i = f\psi_t^e + \frac{X_t^i v_t^e}{D_t^i\mathcal{X}_t^i} + \frac{(X_t^i)^2\psi_t^e}{f} - \frac{(X_t^i v_t^i\cos(\Theta_t^i))}{D_t^i\mathcal{X}_t^i} + \frac{(fv_t^i\sin(\Theta_t^i))}{D_t^i\mathcal{X}_t^i} \tag{36}$$

and

$$\nu_t^i = \frac{Y_t^i(fv_t^e - fv_t^i\cos(\Theta_t^i) + X_t^i\psi_t^e D_t^i\mathcal{X}_t^i)}{fD_t^i\mathcal{X}_t^i}, \tag{37}$$

which gives

$$D_t^i(\mu_t^i) = -\frac{f(X_t^i v_t^e - X_t^i v_t^i\cos(\Theta_t^i) + fv_t^i\sin(\Theta_t^i))}{\mathcal{X}_t^i(\psi_t^e(X_t^i)^2 + \psi_t^e f^2 - \mu_t^i f)} \tag{38}$$

and

$$D_t^i(\nu_t^i) = \frac{Y_t^i f(v_t^e - v_t^i\cos(\Theta_t^i))}{\mathcal{X}_t^i(f\nu_t^i - X_t^i Y_t^i\psi_t^e)}. \tag{39}$$

Since the only unknown information in (38) and (39) is v_t^i and Θ_t^i, we can find the range of D_t^i characterized by $\underline{D}_t^i \leq D_t^i \leq \overline{D}_t^i$, where

$$\underline{D}_t^i = \max\{\min_{\underline{v}^i \leq v_t^i \leq \overline{v}^i, 0 \leq \Theta_t < 2\pi} D_t^i(\mu_t^i), \min_{\underline{v}^i \leq v_t^i \leq \overline{v}^i, 0 \leq \Theta_t < 2\pi} D_t^i(\nu_t^i)\}, \tag{40}$$

$$\overline{D}_t^i = \min\{\max_{\underline{v}^i \leq v_t^i \leq \overline{v}^i, 0 \leq \Theta_t < 2\pi} D_t^i(\mu_t^i), \max_{\underline{v}^i \leq v_t^i \leq \overline{v}^i, 0 \leq \Theta_t < 2\pi} D_t^i(\nu_t^i)\}. \tag{41}$$

We have

$$\frac{\partial D_t^i(\mu_t^i)}{\partial v_t^i} = -\frac{f(-X_t^i \cos(\Theta_t^i) + f \sin(\Theta_t^i))}{\mathcal{X}_t^i(\psi_t^e (X_t^i)^2 + \psi_t^e f^2 - \mu_t^i f)} \tag{42}$$

and

$$\frac{\partial D_t^i(\mu_t^i)}{\partial \Theta_t^i} = -\frac{f(X_t^i v_t^i \sin(\Theta_t^i) + f v_t^i \cos(\Theta_t^i))}{\mathcal{X}_t^i(\psi_t^e (X_t^i)^2 + \psi_t^e f^2 - \mu_t^i f)}. \tag{43}$$

Given the simple form of (42) and (43), we can easily identify the zeros of these equations. We also know that the global optimal solutions of the optimization problems given in lines 3 and 4 in Algorithm 1 occur when $\frac{\partial D_t^i(\mu_t^i)}{\partial v_t^i} = \frac{\partial D_t^i(\mu_t^i)}{\partial \Theta_t^i} = 0$, $v_t^i - \underline{v}_t^i = \frac{\partial D_t^i(\mu_t^i)}{\partial \Theta_t^i} = 0$, or $v_t^i - \overline{v}_t^i = \frac{\partial D_t^i(\mu_t^i)}{\partial \Theta_t^i} = 0$ holds. Therefore, we only need to evaluate a finite number of values to obtain the guaranteed optimal solutions to the optimization problems given in lines 3 and 4 in Algorithm 1. Similarly, we have

$$\frac{\partial D_t^i(\nu_t^i)}{\partial v_t^i} = -\frac{Y_t^i f \cos(\Theta_t^i)}{\mathcal{X}_t^i(f\nu_t^i - X_t^i Y_t^i \psi_t^e)} \tag{44}$$

and

$$\frac{\partial D_t^i(\nu_t^i)}{\partial \Theta_t^i} = \frac{Y_t^i f v_t^i \sin(\Theta_t^i)}{\mathcal{X}_t^i(f\nu_t^i - X_t^i Y_t^i \psi_t^e)} \tag{45}$$

We can also easily identify the zeros of (44) and (45). Similar to the case of $D_t^i(\mu_t^i)$, we only need to evaluate a finite number of values to obtain the guaranteed optimal solutions to the optimization problems given in lines 5 and 6 in Algorithm 1.

We have

$$\dot{\mu}_t^i = \frac{2\nu_t^i \mu_t^i}{Y_t^i} + \frac{f\ddot{y}_t^{\mathrm{v}_i} - X_t^i \ddot{x}_t^{\mathrm{v}_i}}{x_t^{\mathrm{v}_i}} \tag{46}$$

and

$$\dot{\nu}_t^i = \frac{2(\nu_t^i)^2}{Y_t^i} - \frac{\ddot{x}_t^{\mathrm{v}_i}}{x_t^{\mathrm{v}_i}} Y_t^i. \tag{47}$$

To evaluate (46) and (47), we find

$$\ddot{x}_t^{\mathrm{v}_i} = \frac{d\dot{x}_t^{\mathrm{v}_i}}{dv_t^e}\frac{dv_t^e}{dt} + \frac{d\dot{x}_t^{\mathrm{v}_i}}{dv_t^i}\frac{dv_t^i}{dt} + \frac{d\dot{x}_t^{\mathrm{v}_i}}{d\theta_t^e}\frac{d\theta_t^e}{dt} + \frac{d\dot{x}_t^{\mathrm{v}_i}}{d\theta_t^i}\frac{d\theta_t^i}{dt} + \frac{d\dot{x}_t^{\mathrm{v}_i}}{dy_t^{\mathrm{v}_i}}\frac{dy_t^{\mathrm{v}_i}}{dt} + \frac{d\dot{x}_t^{\mathrm{v}_i}}{d\psi_t^e}\frac{d\psi_t^e}{dt} \tag{48}$$

$$= -x_t^{\mathrm{v}_i}(\psi_t^e)^2 - 2v_t^i \sin(\theta_t^e - \theta_t^i)\psi_t^e - a_t^e - \dot{\psi}_t^e y_t^{\mathrm{v}_i} + a_t^i \cos(\theta_t^e - \theta_t^i) + \psi_t^i v_t^i \sin(\theta_t^e - \theta_t^i) \tag{49}$$

$$= -x_t^{\mathrm{v}_i}(\psi_t^e)^2 - 2v_t^i \sin(\Theta_t^i)\psi_t^e - a_t^e - \dot{\psi}_t^e y_t^{\mathrm{v}_i} + a_t^i \cos(\Theta_t^i) + \psi_t^i v_t^i \sin(\Theta_t^i) \tag{50}$$

and

$$\ddot{y}_t^{\mathrm{v}_i} = \frac{d\dot{y}_t^{\mathrm{v}_i}}{dx_t^{\mathrm{v}_i}}\frac{dx_t^{\mathrm{v}_i}}{dt} + \frac{d\dot{y}_t^{\mathrm{v}_i}}{d\psi_t^e}\frac{d\psi_t^e}{dt} + \frac{d\dot{y}_t^{\mathrm{v}_i}}{dv_t^i}\frac{dv_t^i}{dt} + \frac{d\dot{y}_t^{\mathrm{v}_i}}{d\theta_t^e}\frac{d\theta_t^e}{dt} + \frac{d\dot{y}_t^{\mathrm{v}_i}}{d\theta_t^i}\frac{d\theta_t^i}{dt} \tag{51}$$

$$= \psi_t^e x_t^{\mathrm{v}_i} - \psi_t^e v_t^e - (\psi_t^e)^2 y_t^{\mathrm{v}_i} + a_t^i \sin(\theta_t^e - \theta_t^i) + 2\psi_t^e v_t^i \cos(\theta_t^e - \theta_t^i) - \psi_t^i v_t^i \cos(\theta_t^e - \theta_t^i) \tag{52}$$

$$= \psi_t^e x_t^{\mathrm{v}_i} - \psi_t^e v_t^e - (\psi_t^e)^2 y_t^{\mathrm{v}_i} + a_t^i \sin(\Theta_t^i) + 2\psi_t^e v_t^i \cos(\Theta_t^i) - \psi_t^i v_t^i \cos(\Theta_t^i). \tag{53}$$

Therefore, plugging in (34) and (35), we have

$$\dot{\mu}_t^i = \phi_t^e f + \frac{2\mu_t^i \nu_t^i}{Y_t^i} + \frac{(X_t^i)^2 \phi_t^e}{f} + \frac{X_t^i a_t^e}{D_t^i \mathcal{X}_t^i} + \frac{a_t^i f \sin(\Theta_t^i)}{D_t^i \mathcal{X}_t^i} - \frac{f\psi_t^e v_t^e}{D_t^i \mathcal{X}_t^i} - \frac{X_t^i a_t^i \cos(\Theta_t^i)}{D_t^i \mathcal{X}_t^i} + \frac{2f\psi_t^e v_t^i \cos(\Theta_t^i)}{D_t^i \mathcal{X}_t^i}$$
$$- \frac{f\psi_t^i v_t^i \cos(\Theta_t^i)}{D_t^i \mathcal{X}_t^i} + \sin(\Theta_t^i)\left(\frac{2X_t^i \psi_t^e v_t^i}{D_t^i \mathcal{X}_t^i} - \frac{X_t^i \psi_t^i v_t^i}{D_t^i \mathcal{X}_t^i}\right) \tag{54}$$

and

$$\dot{\nu}_t^i = Y_t^i(\psi_t^e)^2 + \frac{2Y_t^i v_t^i \sin(\Theta_t^i)\psi_t^e}{D_t^i \mathcal{X}_t^i} + \frac{2(\nu_t^i)^2}{Y_t^i} + \frac{Y_t^i a_t^e}{D_t^i \mathcal{X}_t^i} + \frac{X_t^i Y_t^i \phi_t^e}{f} - \frac{Y_t^i a_t^i \cos(\Theta_t^i)}{D_t^i \mathcal{X}_t^i} - \frac{Y_t^i \psi_t^i v_t^i \sin(\Theta_t^i)}{D_t^i \mathcal{X}_t^i}. \tag{55}$$

We have the following partial derivatives about $\dot{\mu}_t^i$:

$$\frac{\partial \dot{\mu}_t^i}{\partial D_t^i} = -\frac{X_t^i a_t^e - f\psi_t^e v_t^e + \cos(\Theta_t^i)\left(-X_t^i a_t^i + 2f\psi_t^e v_t^i - f\psi_t^i v_t^i\right) + \sin(\Theta_t^i)\left(2X_t^i \psi_t^e v_t^i - X_t^i \psi_t^i v_t^i + a_t^i f\right)}{(D_t^i)^2 \mathcal{X}_t^i}, \tag{56}$$

$$\frac{\partial \dot{\mu}_t^i}{\partial v_t^i} = \frac{\left(f\cos(\Theta_t^i) + X_t^i \sin(\Theta_t^i)\right)\left(2\psi_t^e - \psi_t^i\right)}{D_t^i \mathcal{X}_t^i}, \tag{57}$$

$$\frac{\partial \dot{\mu}_t^i}{\partial \Theta_t^i} = \frac{X_t^i a_t^i \sin(\Theta_t^i) + f a_t^i \cos(\Theta_t^i) + 2X_t^i \psi_t^e v_t^i \cos(\Theta_t^i) - X_t^i \psi_t^i v_t^i \cos(\Theta_t^i) - 2f\psi_t^e v_t^i \sin(\Theta_t^i) + f\psi_t^i v_t^i \sin(\Theta_t^i)}{D_t^i \mathcal{X}_t^i}, \tag{58}$$

$$\frac{\partial \dot{\mu}_t^i}{\partial \psi_t^i} = -\frac{v_t^i(f\cos(\Theta_t^i) + X_t^i \sin(\Theta_t^i))}{D_t^i \mathcal{X}_t^i}, \tag{59}$$

$$\frac{\partial \dot{\mu}_t^i}{\partial a_t^i} = \frac{f\sin(\Theta_t^i) - X_t^i \cos(\Theta_t^i)}{D_t^i \mathcal{X}_t^i}. \tag{60}$$

There are 2 cases where the optimization problems given in lines 10 and 17 in Algorithm 1 attain their optimal solutions. The first case is when $\frac{\partial \dot{\mu}_t^i}{\partial D_t^i} = \frac{\partial \dot{\mu}_t^i}{\partial v_t^i} = \frac{\partial \dot{\mu}_t^i}{\partial \Theta_t^i} = \frac{\partial \dot{\mu}_t^i}{\partial \psi_t^i} = \frac{\partial \dot{\mu}_t^i}{\partial a_t^i} = 0$. Although the zeros of (56) and (58) are difficult to find, we can easily identify the zeros in (57), (59), and (60). These zeros are sufficient in limiting the search area to a finite number of values. The second case is when D_t^i, v_t^i, ψ_t^i and a_t^i take their maximum or minimum values, and $\frac{\partial \dot{\mu}_t^i}{\partial \Theta_t^i} = 0$. For each specific value that D_t^i, v_t^i, ψ_t^i and a_t^i take, we can easily identify the value of Θ_t^i such that $\frac{\partial \dot{\mu}_t^i}{\partial \Theta_t^i} = 0$. Therefore, the search area is also finite for this case. As a result, we only need to evaluate a finite number of values to obtain the guaranteed optimal solutions to the optimization problems given in lines 10 and 17 in Algorithm 1. Similarly, we have the following partial derivatives about $\dot{\nu}_t^i$:

$$\frac{\partial \dot{\nu}_t^i}{\partial D_t^i} = -\frac{Y_t^i(a_t^e - a_t^i\cos(\Theta_t^i) + 2\psi_t^e v_t^i \sin(\Theta_t^i) - \psi_t^i v_t^i \sin(\Theta_t^i))}{(D_t^i)^2 \mathcal{X}_t^i}, \tag{61}$$

$$\frac{\partial \dot{\nu}_t^i}{\partial v_t^i} = \frac{Y_t^i \sin(\Theta_t^i)(2\psi_t^e - \psi_t^i)}{D_t^i \mathcal{X}_t^i}, \tag{62}$$

$$\frac{\partial \dot{\nu}_t^i}{\partial \Theta_t^i} = \frac{Y_t^i(a_t^i \sin(\Theta_t^i) + 2\psi_t^e v_t^i \cos(\Theta_t^i) - \psi_t^i v_t^i \cos(\Theta_t^i))}{D_t^i \mathcal{X}_t^i}, \tag{63}$$

$$\frac{\partial \dot{\nu}_t^i}{\partial \psi_t^i} = -\frac{Y_t^i v_t^i \sin(\Theta_t^i)}{D_t^i \mathcal{X}_t^i}, \tag{64}$$

$$\frac{\partial \dot{\nu}_t^i}{\partial a_t^i} = -\frac{Y_t^i \cos(\Theta_t^i)}{D_t^i \mathcal{X}_t^i}. \tag{65}$$

Similar to the case of $\dot{\mu}_t^i$, we only need to evaluate a finite number of values to obtain the guaranteed optimal solutions to the optimization problems given in lines 27 and 34 in Algorithm 1. ■

Proof (Theorem 1). We first show that, given that $\mu_0^i > \mu_b^i$, and $\dot{\mu}_t^i > 0$ whenever $\mu_0^i \leq \mu_b^i$, we have that $\mu_t^i > \mu_b^i$ for all time $t \geq 0$. We prove that using contradiction. Suppose that $\mu_0^i > \mu_b^i$, and $\dot{\mu}_t^i > 0$ whenever $\mu_0^i \leq \mu_b^i$, and there exists t_m such that $\mu_{t_m}^i < \mu_b^i$. We have

$$\mu_t^i = f\frac{\dot{y}_t^{\mathrm{v}_i} x_t^{\mathrm{v}_i} - y_t^{\mathrm{v}_i} \dot{x}_t^{\mathrm{v}_i}}{(x_t^{\mathrm{v}_i})^2}. \tag{66}$$

Since $x_t^{\mathrm{v}_i} > 0$, and $x_t^{\mathrm{v}_i}$ and $y_t^{\mathrm{v}_i}$, as well as their time derivatives, are continuous in t for realistic objects, we have that μ_t^i is uniformly continuous in t. From intermediate value theorem, there exist $\tau \in (0, t_m)$ such that $\mu_\tau^i = \mu_b^i$. Let τ_m be the largest such τ, where we have $\mu_t^i < \mu_b^i$ for all $t \in (\tau_m, t_m]$. Since $\dot{\mu}_{\tau_m}^i > 0$, we know that there exists $\epsilon > 0$ such that $\mu_{\tau_m + t_\epsilon}^i > \mu_b^i$ for all $t_\epsilon \in (0, \epsilon)$. This contradicts the statement that $\mu_t^i < \mu_b^i$ for all $t \in (\tau_m, t_m]$. Line 17 of Algorithm 1 ensures that $\dot{\mu}_t^i > \underline{\dot{\mu}}_t^i > 0$ whenever $\mu_0^i \leq \mu_b^i$. Therefore, $\mu_t^i > \mu_b^i$ for all time $t \geq 0$. The proof for $\mu_t^i < \mu_b^i$ is analogous and is not explained in detail here. ■

Proof (Theorem 2). We first show that, given that $\nu_0^i > \nu_b^i$, and $\dot{\nu}_t^i > 0$ whenever $\nu_0^i \leq \nu_b^i$, we have that $\nu_t^i > \nu_b^i$ for all time $t \geq 0$. We prove that using contradiction. We have

$$\nu_t^i = f\frac{\dot{z}_t^{\mathrm{v}_i} x_t^{\mathrm{v}_i} - z_t^{\mathrm{v}_i} \dot{x}_t^{\mathrm{v}_i}}{(x_t^{\mathrm{v}_i})^2}. \tag{67}$$

Since $x_t^{\mathrm{v}_i} > 0$, and $x_t^{\mathrm{v}_i}$ and $z_t^{\mathrm{v}_i}$, as well as their time derivatives, are continuous in t for realistic objects, we have that ν_t^i is uniformly continuous in t. Following the same step as the proof for Theorem 1 with line 34 in Algorithm 1, we have $\nu_t^i > \nu_b^i$ for all time $t \geq 0$. The proof for $\nu_t^i < \nu_b^i$ is analogous and is not explained in detail here. ■

Proof (Theorem 3). Consider the Lyapunov function given by

$$V(X_t^i) = \frac{1}{2}(X_t^i)^2, \tag{68}$$

where

$$\dot{V}(X_t^i) = X_t^i \dot{X}_t^i = X_t^i \mu_t^i < 0 \tag{69}$$

for all $X_t^i \neq 0$ ensures that $X_t^i \to 0$. Since $x_t^{\mathrm{v}_i} > 0$, we have that for finite $x_t^{\mathrm{v}_i}$, $X_t^i \to 0$ implies $y_t^{\mathrm{v}_i} \to 0$. ■

Proof (Lemma 2). From (3), we have

$$x^{\mathrm{v}} = \frac{f z^{\mathrm{v}}}{Y} \tag{70}$$

$$y^{\mathrm{v}} = \frac{X z^{\mathrm{v}}}{Y}. \tag{71}$$

Therefore,

$$\min_{\overline{h} \leq h \leq \underline{h}} r^{\mathrm{v}}(\frac{Xh}{Y}, f\frac{h}{Y}) \geq 0 \tag{72}$$

ensures that for all possible z^{v}, *i.e.,* $\underline{h} \geq z^{\mathrm{v}} \geq \overline{h}$, we have

$$r^{\mathrm{v}}(x^{\mathrm{v}}, y^{\mathrm{v}}) \geq 0, \tag{73}$$

which indicates (19). ■

Proof (Theorem 4). From (3), we have

$$x_t^{\mathrm{v}_i} = \frac{f z_t^{\mathrm{v}_i}}{Y_t^i} \tag{74}$$

$$y_t^{\mathrm{v}_i} = \frac{X_t^i z_t^{\mathrm{v}_i}}{Y_t^i}. \tag{75}$$

Plugging into (20), we have

$$(X_t^i)^2 (z_t^{\mathrm{v}_i})^2 - (Y_t^i)^2 r^2 + (z_t^{\mathrm{v}_i})^2 f^2 \geq 0 \tag{76}$$

implies

$$(x^{\mathrm{v}_i})^2 + (y^{\mathrm{v}_i})^2 - r^2 \geq 0, \tag{77}$$

Since $z_t^{\mathrm{v}_i} \geq \overline{h}$, we have

$$(X_t^i)^2 \overline{h}^2 - (Y_t^i)^2 r^2 + \overline{h}^2 f^2 \geq (X_t^i)^2 (z_t^{\mathrm{v}_i})^2 - (Y_t^i)^2 r^2 + (z_t^{\mathrm{v}_i})^2 f^2, \tag{78}$$

which gives that

$$(X_t^i)^2 \overline{h}^2 - (Y_t^i)^2 r^2 + \overline{h}^2 f^2 := r^o \geq 0 \tag{79}$$

implies (77). We have

$$\dot{r}^o = \overline{h}^2 X_t^i \dot{X}_t^i - r^2 Y_t^i \dot{Y}_t^i = \overline{h}^2 X_t^i \mu_t^i - r^2 Y_t^i \nu_t^i. \tag{80}$$

From Nagumo's theorem [27], assuming $z_t^{\mathrm{v}_i} \geq \overline{h}$ for all time t and $\{x_0^{\mathrm{v}_i}, y_0^{\mathrm{v}_i}\} \in \mathcal{C}^v$, given that

$$\dot{r}^o > 0 \tag{81}$$

is satisfied whenever

$$r^o \leq 0, \tag{82}$$

we have

$$r^o \geq 0 \tag{83}$$

at all time $t > 0$, which implies $\{x_t^{\mathrm{v}_i}, y_t^{\mathrm{v}_i}\} \in \mathcal{C}^{\mathrm{v}}$ at all time $t > 0$. ■