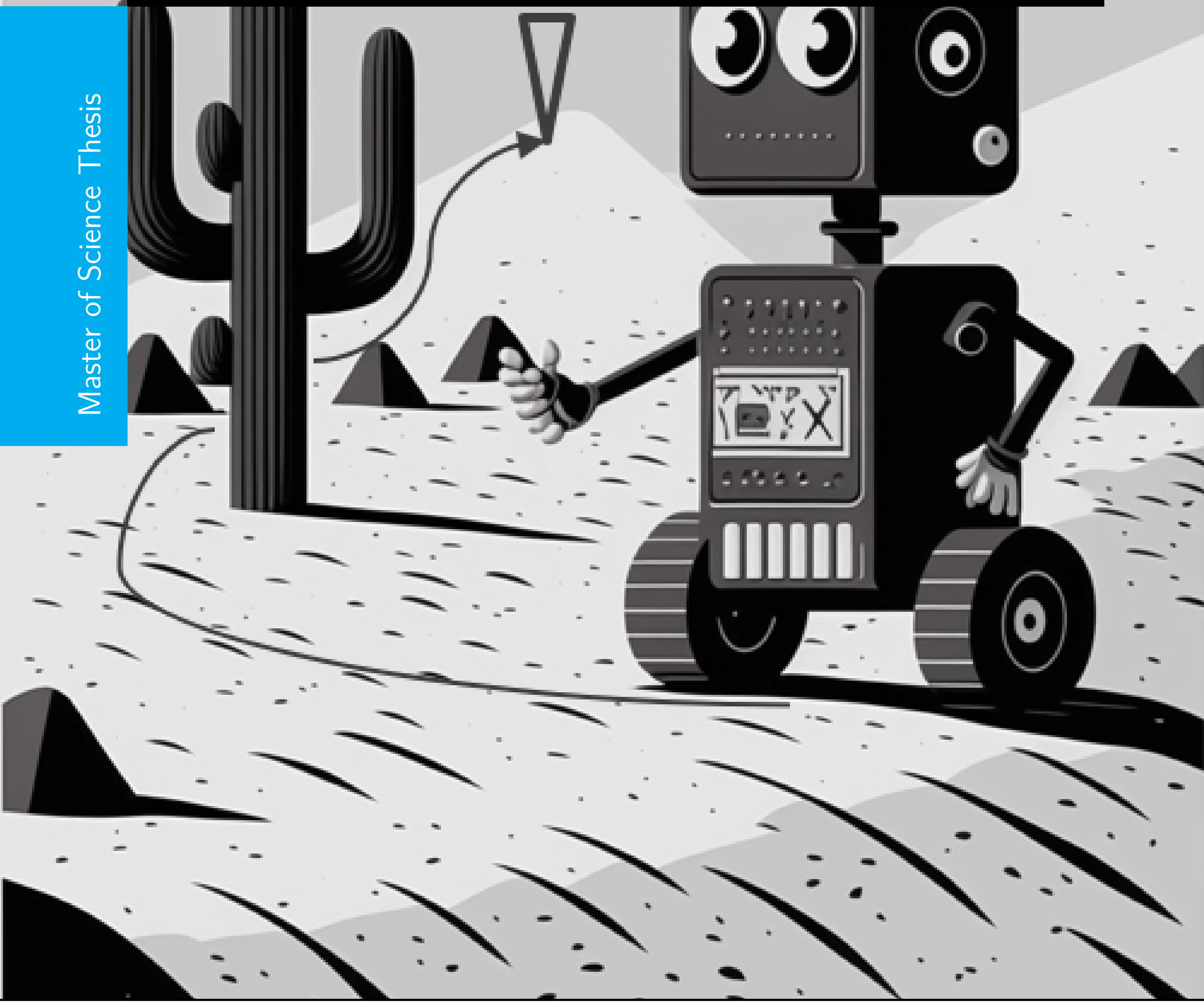
Hierarchical Active Inference Control for a nonholonomic mobile robot

B.P. Benist

Master of Science Thesis



Delft University of Technology

Delft Center for Systems and Control

Hierarchical Active Inference Control for a nonholonomic mobile robot

Master of Science Thesis

For the degrees of Master of Science in Systems and Control and Master of Science in Robotics at Delft University of Technology

B.P. Benist

May 10, 2023

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology

The work in this thesis was supported by Avular. Their cooperation is hereby gratefully acknowledged.

Delft University of Technology
Department of
Delft Center for Systems and Control (dcsc)

The undersigned hereby certify that they have read and recommend to the Faculty of Mechanical, Maritime and Materials Engineering (3mE) for acceptance a thesis entitled

Hierarchical Active Inference Control for a nonholonomic mobile robot

by

B.P. Benist

in partial fulfillment of the requirements for the degree of

Master of Science Systems and Control

Dated: May 10, 2023

Supervisor(s):

Prof.dr.ir. M. Wisse

Dr. P. Mohajerin Esfahani

Dr.ir. A. Andriën

Reader(s):

Dr.ir. J. Sijs

Delft University of Technology
Department of
Cognivitive Robotics (COR)

The undersigned hereby certify that they have read and recommend to the Faculty of Mechanical, Maritime and Materials Engineering (3mE) for acceptance a thesis entitled

Hierarchical Active Inference Control for a nonholonomic mobile robot

by

B.P. Benist

in partial fulfillment of the requirements for the degree of

Master of Science Robotics

Dated: May 10, 2023

Supervisor(s): Prof.dr.ir. M. Wisse

Dr. P. Mohajerin Esfahani

Dr.ir. A. Andriën

Reader(s): Dr.ir. J. Sijs

Abstract

This master thesis introduces Hierarchical Active Inference Control (HAIC) as a control method for nonholonomic systems. This method only requires tuning of a minimal number of hyperparameters and has a relative low computation load. HAIC is based on recent research done in the application of the neuroscientific theory of Active Inference for robot control. The hierarchical aspect of this method introduces two layers of Active Inference Control (AIC)[1] and the introduction of an additional decision hyperparameter. The top layer AIC is designed such that it takes the nonholonomic constraint into account. This causes HAIC to be able to control a nonholonomic mobile robot to a three-dimensional reference state in comparison to regular AIC which is unable to. The selection of the introduced hyperparameter gives a trade-off between robustness against noise present in the measurements and possible error in the y-dimension of the local reference frame.

Convergence of a two-wheeled differential-drive mobile robot to a reference state using HAIC is demonstrated both by simulation and physical experiment. Simulations are done for a large range of different starting states. These simulations show the ability of HAIC to converge the mobile robot for any starting state. It also confirms that a higher value for the decision parameter causes a larger error. Another group of simulations is done investigating the new hyperparameter introduced with HAIC. These simulations check the ability to converge the system when different noise levels are present in the measurements. It is concluded that tuning of the newly introduced hyperparameter needs to accommodate the noise present in the measurements. Where the higher the noise, the higher the value of the hyperparameter is needed for convergence.

A physical experiment is done where the robot is controlled to five different reference states. Data obtained from this experiment shows the ability of HAIC to control a physical nonholonomic system. Like the simulations, it is also shown that a correct value is required to set the newly introduced hyperparameter. Additionally, the physical experiment shows that HAIC is not robust when disturbances can cause sudden large value changes in the measurements.

Table of Contents

Chapter 1

Introduction

An increase in the use of automated mobile robots for inspection and maintenance has been forecasted for the coming years [2, 3]. Often times, the mobile robot directly substitutes the tasks done by a human. These mobile robots therefore consist of unmanned ground vehicles (UGVs) making use of wheeled or legged locomotion [4, 5, 6], and unmanned air vehicles (UAVs) making use of rotors [7, 8].

A drawback for most UAVs is their high-energy consumption in combination with their limited maximum take-off weight, which results in limited operational time before refuelling or recharging [9, 10]. UGVs making use of legged locomotion have an increased complexity with their versatile motions and increased energy consumption compared to wheeled robots [11, 12]. It is therefore preferred to use wheeled robots when these can be applied in environments requiring inspection and maintenance.

An important aspect of automated mobile robots for inspection and maintenance is their capability of navigating through the environment. This navigation consists of top-level navigation such as obstacle avoidance, localization, and trajectory planning, but also low-level navigation such as reference tracking.

The use of a simple wheeled mobile robot (WMR) allows for robustness against rougher terrain, dirt build-up, and hardware failure. A simple WMR makes use of redundant fixed actuated wheels instead of more complex omnidirectional actuated wheels. However, it becomes more challenging to create a trajectory such that these simple WMRs can reach a desired location. The reason for this challenge is because of the nonholonomic characteristic of simple WMRs. A nonholonomic system has a constraint on its velocities but not on its positions. This nonholonomic constraint makes it impossible to determine a time-invariant feedback control law to control the system to a set point [13, 14]. An example of a simple WMR with a nonholonomic constraint is that of a two-wheeled differential-drive mobile robot. The nonholonomic constraint of this mobile robot is that it cannot gain a non-zero velocity in the direction perpendicular to the direction it is facing.

1-1 Research Gap

To still successfully control a nonholonomic system to a reference point, different methods are used to determine a control input. Successful control methods are sliding mode control (SMC) [15, 16], adaptive control using Back-stepping [17, 18], Lyapunov based control [19, 20], and vector-field control[21]. These methods require a relatively low online computation effort but do require non-trivial controller design, Lyapunov function selection, and vector-field creation, respectively.

On the other hand, a lot of research has been done in the application of Model Predictive Control (MPC) for the control of simple WMRs [22]. MPC only requires the selection of a handful of hyperparameters to have the ability to make the system converge to a reference state. However, an issue with MPC for control is the relatively high online computation required for the control input. In particular, in the case of nonholonomic WMRs, MPC needs to solve a non-linear optimization problem with terminal constraints. Recent research did find a possibility to create an MPC controller without a terminal constraint [23]. However, this control method still requires a tailor-made stage cost and has to solve an unconstrained non-linear optimization problem. As of now, methods for controlling nonholonomic systems require many pre-selected parameters or have to solve online non-linear optimization problems.

Active Inference Control (AIC) is a control method that has both the characteristics of a low number of required pre-selected parameters and makes use of simple online computational operations. AIC has the advantage of having a limited amount of preselected parameters which require minimal tuning because it is less sensitive to variations [1]. The operations needed for every control iteration of AIC are a handful of linear operations which scale directly with the dimension of states the system operates in. This is much less computationally heavy in comparison to the non-linear optimizations of multidimensional variables required for MPC every control iteration. For that reason, this thesis further researches AIC for the control of nonholonomic mobile robots.

AIC is based on the neuroscientific theory of Active Inference [24]. The theory is based on the idea that action and perception are used for the cognitive process of the brain to minimize "surprise". Here, the surprise is linked to a value function referred to as the free-energy. The value of the free-energy would be high when an expected perception is not met, based on which an action can be deduced to obtain the expected perception.

Active Inference is related to the control within a human body. The theory of Active Inference models the behaviour of the control of the human body on different "layers". For example, bottom layer control would be the temperature regulation of an organ, while top layer control would be the movement of the hand to grab an object. The bottom layer refers to processes that are highly automated and require minimal observation processing. The top refers to high-level control such as complex motions which require the processing of a high number of observations and inferring multidimensional control inputs. This multi-layer use of Active Inference is called Hierarchical Active Inference (HAI) [25].

Relating to robot control, in the past years multiple applications have been researched that apply the free-energy minimization to obtain control inputs for robot control [26, 27]. Control applications already exist for robot manipulators [1, 28, 29] making use of AIC. Research has also been done on HAI for control, where a top layer using Active Inference infers a goal state for lower-level controllers [30, 31]. These methods make use of HAI as a guideline for the structure of both high-level path planning and low-level control but do not use AIC to determine the control inputs. Instead, often time make use of Proportional Integral Derivative (PID) control, Neural Network (NN), or other methods for sensor processing, and actuator control. To the best of the author's knowledge, HAI and AIC have yet to be combined in general, and on WMRs in particular.

In this Master Thesis, the first step in the direction of the development of Hierarchical Active Inference Control (HAIC) is taken. HAIC combines the structure of HAI with the control aspect of AIC. First, the need for HAIC for a nonholonomic system is shown. After which, through simulation and physical experiments, it is shown that HAIC can steer a nonholonomic system to a reference state. This shows that HAIC is a viable control method when manual selection of parameters and high computation effort are undesirable.

1-2 Thesis Outline

This master thesis researches the capabilities of HAIC for the control of a nonholonomic system. The research question defining this master thesis is: "Can Hierarchical Active Inference Control (HAIC) control a two-wheeled differential-drive mobile robot to a three-dimensional reference state?"

To answer the research question, sub-questions are defined which will be answered throughout this thesis:

1. What are the limitations of AIC for controlling a nonholonomic mobile robot to a three-dimensional reference state?
2. How can AIC and HAI be combined such that it can successfully control a two-wheeled mobile robot to a three-dimensional reference state?
3. Which parameters have to be set in addition to AIC, and what influences their values, such that convergence can be obtained when using HAIC for the control of a two-wheeled mobile robot?
4. How robust is the proposed HAIC controller to noise and sudden value changes present in the measurements?

The reason for sub-question one is to show that a more complex version of AIC is required to control a nonholonomic system. Sub-question two is defined to elaborate on the workings of HAIC and how it is designed such that it can control a nonholonomic system. The answer to question three will verify whether HAIC indeed has a minimal number of hyperparameters and what indication can be used for setting correct values for these parameters. Sub-question four is used to obtain an indication of the robustness of HAIC when noise and other disturbances are present in the measurements.

This Master Thesis is organized by first stating the problem formulation of controlling a nonholonomic system in a three-dimensional space in chapter 2. After which, in chapter 3, the limitation of AIC is shown and how HAIC is designed such that it can control a nonholonomic system. In chapter 4 it is shown in simulation that HAIC can control the system from different starting states and the effect of noise in the observations is shown. HAIC is also tested on a real robot for which results are discussed in chapter 5. In the last chapter, conclusions and future work are summarized.

Chapter 2

Nonholonomic system

In this chapter, the formulation of the nonholonomic system is given. The kinematic model for the system used for both the model of the wheeled mobile robot (WMR) and the control of the system is stated. Assumptions about the environment and dynamics of the system and environment are clarified.

2-1 Kinematic Model

Throughout this master thesis, the nonholonomic system of a two-wheeled differential-drive mobile robot will be controlled. The robot shown in Figure 2-1 operates in a 2D workspace, which consists of positions x and y, and orientation θ. These are combined in the vector z as shown in (2-1).

$$z^F(t) = \begin{bmatrix} x^F(t) \\ y^F(t) \\ \theta^F(t) \end{bmatrix} \tag{2-1}$$

The superscript F indicates the frames according to Figure 2-1, of which there is the global frame (g), and the reference frame (r). Neither of these frames are inertial frames since the global and reference frame do not move. The kinematics of the mobile robot in continuous-time are written according to the unicycle model as:

$$\dot{z}^F(t) = \begin{bmatrix} \dot{x}^F(t) \\ \dot{y}^F(t) \\ \dot{\theta}^F(t) \end{bmatrix} = \begin{bmatrix} v^F(t)cos(\theta^F(t)) \\ v^F(t)sin(\theta^F(t)) \\ \omega^F(t) \end{bmatrix}, \quad \text{for: } F = g, r \tag{2-2}$$

The forward velocity is indicated by $v(t)$ and $\omega(t)$ indicates the angular velocity. These two variables are considered the input $u(t)$ of the nonholonomic system throughout this thesis. It is assumed there is no wheel slip present.

The relation between a frame A and a frame B is shown in (2-3). The superscript A and B indicate in which frame the position is given, while the subscript A and B indicate the

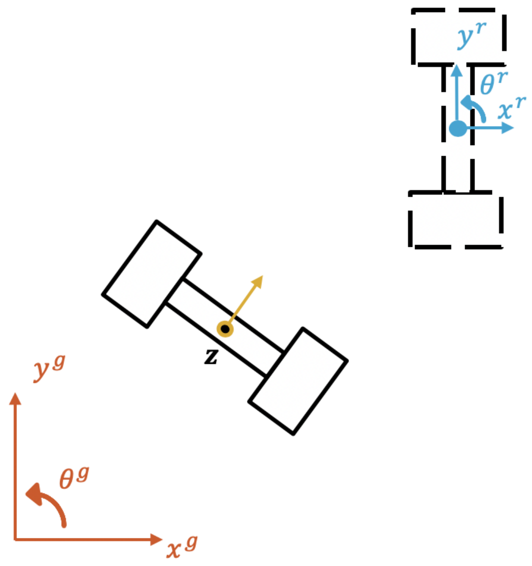


Figure 2-1: The 2D workspace in which the nonholonomic robot operates. The frame indicated with red and the superscript g is the global frame and the frame indicated with blue and the superscript r is the reference frame. The vector z is the position of the center of the robot. The yellow arrow indicates the heading of the robot.

position of the origin of the frame. Note that throughout this thesis, z indicates the position of the robot.

$$\begin{aligned}\begin{bmatrix} x^B \\ y^B \\ \omega^B \end{bmatrix} &= \begin{bmatrix} cos(\theta_B^A) & sin(\theta_B^A) & 0 \\ -sin(\theta_B^A) & cos(\theta_B^A) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^A - x_B^A \\ y^A - y_B^A \\ \theta^A - \theta_B^A \end{bmatrix} \\ \begin{bmatrix} \dot{x}^B \\ \dot{y}^B \\ \omega^B \end{bmatrix} &= \begin{bmatrix} cos(\theta_B^A) & sin(\theta_B^A) & 0 \\ -sin(\theta_B^A) & cos(\theta_B^A) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{x}^A \\ \dot{y}^A \\ \dot{\theta}^A \end{bmatrix}\end{aligned} \tag{2-3}$$

The nonholonomic aspect of a two-wheeled differential-drive system is that it is unable to have a velocity in the direction perpendicular to its heading. A nonholonomic system has a constraint on the velocity but does not have a constraint on its position. For a two-wheeled differential-drive mobile robot in a 2D workspace, the constraint on the velocity is present in the x and y dimension and dependent on its orientation θ. The nonholonomic constraint on the velocity can be written as:

$$\begin{aligned}\dot{y}^b(t) &= \dot{x}^g(t)sin(\theta^g(t)) - \dot{y}^g(t)cos(\theta^g(t)) = 0 \\ \dot{y}^r(t) &= \dot{x}^r(t)sin(\theta^r(t)) - \dot{y}^r(t)cos(\theta^r(t)) = 0\end{aligned} \tag{2-4}$$

2-2 Control Limitations

The nonholonomic constraint of a two-wheeled differential-drive mobile robot causes limitations on possible control techniques. It is impossible to determine a linear time-invariant (LTI) state feedback controller for the control of a nonholonomic system. To be able to control a system with an LTI state feedback controller, the system needs to be able to uphold to Brockett's necessary conditions (Theorem 1)(original from [32] but found via [33]).

For Brockett's necessary condition $z \in \mathbb{R}^n$ and $\dot{z} = f(z, u)$.

Theorem 1 (Brockett's necessary condition)**.** *If a continuously differentiable control law $u = k(z)$ exists rendering z_0 an asymptotically stable equilibrium, then the image of the mapping (z, u) by $f(z, u)$ contains a neighbourhood of 0 in the dimensional space of $\mathbb{R}^n$.*

This does not hold for the unicycle model. The dynamics describing the states z as given in equation (2-2) can be written in the form of:

$$f(z, u) = \begin{bmatrix} cos(\theta) & 0 \\ sin(\theta) & 0 \\ 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} v \\ \omega \end{bmatrix}}_{u} \tag{2-5}$$

Points on the line of $z = [0, a, 0]^T$ where $a \neq 0$ would give the form of:

$$f(z, u) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{2-6}$$

This indicates that there is no value of u, whether based on z or not, that will cause a mapping of (z, u) by $f(z, u)$ to contain the origin of $\mathbb{R}^n$. Proving that there is no control law $u = k(z)$ which will be able to stabilize the system to a reference. The nonholonomic constraint thus causes the system to not uphold to Brockett's condition Theorem 1, which implies that there is no possibility to determine a time-invariant feedback control law to control the system to a set point. For that reason, only time-varying control methods can be used for the control of nonholonomic systems.

2-3 Assumptions And Notations

The robot operates in a three-dimensional environment also known as a two-dimensional workspace. The robot itself has five states: the two position states x and y, an orientation state θ, and the forward and rotational velocities v and ω.

Assumption 1. *A minimal representation of the dynamics of the robot can be given in six states $[x, y, \theta, v, \omega]^T$.*

This master thesis does not focus on obstacle avoidance. For that reason, it is assumed the system operates in an infinite plane in the x and y dimensions without any obstacles present.

Assumption 2. *The mobile robot is operating in an infinite x and y plane free of obstacles.*

Throughout this thesis, control is applied to the states in the reference frame. To make notation easier, the superscript r will not be applied to any equation. Therefore, the following is applicable throughout this thesis:

$$z = z^r, \quad \dot{z} = \dot{z}^r, \quad \begin{bmatrix} x \\ y \\ \theta \end{bmatrix} = \begin{bmatrix} x^r \\ y^r \\ \theta^r \end{bmatrix}, \quad \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \dot{x}^r \\ \dot{y}^r \\ \dot{\theta}^r \end{bmatrix} \tag{2-7}$$

The three position states are known and can be used by the controller. The control input is known at all times. The velocities are calculated according to equation (2-2) and make use of the three known states and the known input. The dynamics of the model uphold to the unicycle model in equation (2-2) and the nonholonomic constraint in equation (2-4).

Assumption 3. *The states of x, y, and θ are directly measurable, and the values of the control inputs v and ω are known.*

Chapter 3

Active Inference Control

In this chapter Active Inference Control (AIC) is explained. The method of AIC in this thesis is based on previous research [1, 28]. First, the main principle of the Active Inference framework is discussed, relating the belief μ to the hidden states z of a system. After which the determination of the control input is introduced, which explains how AIC works. The application of AIC on a nonholonomic system is shown, in which the reason for its inability to converge to the origin from varying starting states is explained. Based on this inability, AIC is adapted to Hierarchical Active Inference Control (HAIC) such that it can converge to the origin.

3-1 Active Inference

Active Inference considers a dynamic system. The dynamics of this system are described with hidden states z. These hidden states are not directly available to the system, but information about them is contained within the available observable states o. The aim is thus to infer the hidden states z from the observable states o.

Free-Energy

The observable and hidden states are represented as a probability distribution which are notated as $p(o)$ and $p(z)$ respectively. The distribution of the hidden states z given the observations o is notated as $p(z|o)$. One can use Bayes' rule to find $p(z|o)$ via:

$$p(z|o) = \frac{p(o|z)p(z)}{p(o)} \tag{3-1}$$

An issue with Bayes' rule is obtaining $p(o)$, which requires the calculation of $\int p(o|z)p(o)do$ for every new obtained observation [34]. For that reason, a variational Bayes approach with a probability distribution $q(z)$ is used to approximate $p(z|o)$. The distribution $q(z)$ is a simpler form of distribution compared to that of $p(z|o)$.

This approximation is obtained using a Kullback-Leibler (KL) divergence between $q(z)$ and $p(z|o)$. The KL divergence of the two distributions represents the difference between the distributions. By minimizing this difference, the distribution $q(z)$ will approximate $p(z|o)$. The KL divergence is given as:

$$\begin{aligned} D_{KL}(q(z)||p(z|o)) &= \int q(z)\ \ln\frac{q(z)}{p(z|o)}dz \qquad p(z|o) = \frac{p(z,o)}{p(o)} \\ &= \int q(z)\ \ln\frac{q(z)}{p(z,o)}dz + \ \ln\ p(o) \\ &= \mathcal{F} + \ln\ p(o) \end{aligned} \tag{3-2}$$

The $\mathcal{F}$ term can be minimized with respect to z to obtain the approximation using variational Bayes. Note that ln $p(o)$ cannot be minimized, only terms containing z are minimized. This results in the system inferring z from o. The $\mathcal{F}$ term is called the variational free-energy [34].

An adaptation of the free energy can be used when choosing $q(z)$ to be a normalized Gaussian distribution with mean μ. A simplified expression for $\mathcal{F}$ can be obtained by having the Gaussian distribution sharply peak at μ, assuming ln (z, o) is a smooth function, and stating that σ^2 of $q(z)$ is already optimized to minimize $\mathcal{F}$ [34]. The final equation is shown in Equation 3-3, all the steps done to reach this simplification can be found in the Appendix in section A-1.

$$\begin{aligned} \mathcal{F} &= \int q(z)\ \ln\frac{q(z)}{p(z,o)}dz \\ &= \underbrace{\int q(z)\ \ln\ q(z)dz}_{\approx 0} - \underbrace{\int q(z)\ \ln\ p(z,o)dz}_{\approx\ \ln\ p(\mu,o)} \\ \mathcal{F} &\approx -\ln\ p(\mu,o) \end{aligned} \tag{3-3}$$

As before, observations o are obtained and cannot be used to minimize, thus minimization of $\mathcal{F}$ becomes dependent on the parameter μ.

The variable μ estimating z is often called the “belief”. This comes from the roots of Active Inference which lay in cognitive neuroscience. Throughout this thesis, μ is referred to as a belief instead of an estimation.

Generative Models

Now a relation between the observations o and the beliefs μ have to be defined. This is done by stating that o has a mean of $g(\mu)$ and a variance of Σ_o. The relation can be expressed as:

$$p(o|\mu) = \mathcal{N}(g(\mu), \Sigma_o) \tag{3-4}$$

Here $\mathcal{N}(g(\mu), \Sigma_o)$ represents a Gaussian distribution and Σ_o represents the noise present in the observations. The variance Σ_o is a square matrix with a number of columns and rows equal to the number of dimensions in the observations. The matrix also represents the confidence of the information about μ contained in each observation o. Where a high value of Σ_o equals a low confidence in the information of μ contained in o.

The relation for a single observation is written in the form of:

$$o = g(\mu) + w_o, \quad w_o \sim \mathcal{N}(0, \Sigma_o) \tag{3-5}$$

Where w_o is additive Gaussian noise with w_o being a value from a distribution $\mathcal{N}(0, \Sigma_o)$.

As stated earlier, Active Inference is applied to a dynamical system. Therefore, the belief μ has its dynamics. These dynamics are represented as:

$$p(\frac{d}{dt}\mu|\mu) = p(\mu'|\mu) = \mathcal{N}(f(\mu), \Sigma_\mu) \tag{3-6}$$

Here, $\mathcal{N}(f(\mu), \Sigma_\mu)$ is a similar Gaussian distribution as $\mathcal{N}(g(\mu), \Sigma_o)$ but Σ_μ represents the process noise. This matrix also represents the confidence in the system acting according to the dynamics defined by $f(\mu)$.

The dynamics of a single belief is written in the form of:

$$\mu' = f(\mu) + w_\mu \tag{3-7}$$

Where w_μ is additive Gaussian noise $w_\mu \sim (0, \Sigma_\mu)$.

Often times, the variances of the noise present in the observations and processes are not known. Relating these to the confidence of an observation or a process allows for the selection of the parameter without knowing the true value. Accounting a low value to Σ_o would be stating that the observations contain an accurate value for the hidden states. Accounting a high value would be stating that the confidence in the observations to contain a correct value of μ is low. This is of importance later when these variables have to be defined when applying Active Inference to a system.

The relation of μ and o, and the dynamics of μ are defined as generative models. Consisting of the generative model of the sensory data for the relation of o and μ, and the generative model of the state dynamics for the evolution of μ.

Generalized Motions

A generalized motion of the system is used to obtain a more accurate representation of the hidden states and the dynamics of the hidden states. This generalized motion consists of higher order generative models of the state dynamics and their correlating generative model of the sensory data.

As stated in equation (3-6), the generative model of the dynamics is represented by μ'. This μ' can have its own generative model of its state dynamics, which is represented by μ''. This relation can be described for higher order derivatives of a belief as:

$$\begin{aligned} \mu' &= f(\mu) + w_\mu, & w_\mu &\sim \mathcal{N}(0, \Sigma_\mu) \\ \mu'' &= \underbrace{\frac{\partial f}{\partial \mu}\mu'}_{f'(\mu')} + w'_\mu, & w'_\mu &\sim \mathcal{N}(0, \Sigma_{\mu'}) \\ \mu^{(i+1)} &= f^{(i)}(\mu^{(i)}) + w^{(i)}_\mu, & w^{(i)}_\mu &\sim \mathcal{N}(0, \Sigma_{\mu^{(i)}}) \end{aligned} \tag{3-8}$$

The generalized motions of the beliefs are indicated as $\tilde{\mu} = [\mu, \mu', \cdots, \mu^{n_d+1}]$. The correlated observations are $\tilde{o} = [o, o', \cdots, o^{n_d}]$ where n_d is the order of the highest observable derivative. The relationship of higher order observations and beliefs are defined, similar to the zeroth order belief, as:

$$
\begin{aligned}
o' &= \underbrace{\frac{\partial g}{\partial \mu}\mu'}_{g'(\mu')} + w_{o'}, \qquad && w_{o'} \sim \mathcal{N}(0, \Sigma_{o'}) \\
o^{(i)} &= g^{(i)}(\mu^{(i)}) + w_{o^{(i)}}, && w_{o^{(i)}} \sim \mathcal{N}(0, \Sigma_{o^{(i)}})
\end{aligned}
\tag{3-9}
$$

The Free-Energy Equation

The generalized motions consisting of the generative models can be used in the free-energy term $\mathcal{F}$. The expectation of $p(\tilde{\mu}, \tilde{o})$ can be given using the generalized motions.

$$
\mathcal{F} = -\ln p(\tilde{\mu}, \tilde{o}) = -\ln \left(\prod_{i=0}^{n_d} p(o^{(i)}|\mu^{(i)}) p(\mu^{(i+1)}|\mu^{(i)}) \right) \tag{3-10}
$$

As stated earlier, the distributions of $p(o^{(i)}|\mu^{(i)})$ and $p(\mu^{(i+1)}|\mu^{(i)})$ are assumed to be normal Gaussian distributions and are therefore written as:

$$
\begin{aligned}
p(o^{(i)}|\mu^{(i)}) &= \frac{1}{|\Sigma_{o^{(i)}}| \sqrt[n_o]{2\pi}} e^{\left(-\frac{1}{2}(o^{(i)} - g^{(i)}(\mu^{(i)}))^T \Sigma_{o^{(i)}}^{-1} (o^{(i)} - g^{(i)}(\mu^{(i)}))\right)} \\
p(\mu^{(i+1)}|\mu^{(i)}) &= \frac{1}{|\Sigma_{\mu^{(i)}}| \sqrt[n_\mu]{2\pi}} e^{\left(-\frac{1}{2}(\mu^{(i+1)} - f^{(i)}(\mu^{(i)}))^T \Sigma_{\mu^{(i)}}^{-1} (\mu^{(i+1)} - f^{(i)}(\mu^{(i)}))\right)}
\end{aligned}
\tag{3-11}
$$

The variables n_o and n_μ are equal to the dimension of o and μ respectively.

Using the representation of a Gaussian distribution, we can rewrite the variational free energy equation.

$$
\begin{aligned}
\mathcal{F} =& -\sum_{i=0}^{n_d} \left(\ln p(o^{(i)}|\mu^{(i)}) + \ln p(\mu^{(i+1)}|\mu^{(i)}) \right) \\
=& \frac{1}{2} \sum_{i=0}^{n_d} \Big((o^{(i)} - g^{(i)}(\mu^{(i)}))^T \Sigma_{o^{(i)}}^{-1} (o^{(i)} - g^{(i)}(\mu^{(i)})) \\
& + (\mu^{(i+1)} - f^{(i)}(\mu^{(i)}))^T \Sigma_{\mu^{(i)}}^{-1} (\mu^{(i+1)} - f^{(i)}(\mu^{(i)})) \Big) + K
\end{aligned}
\tag{3-12}
$$

The variable K is a combination of terms that cannot be minimized, since it does not contain any order of μ. The $\mathcal{F}$ term is further reduced to a term that can be minimized to obtain a correct approximation.

The belief μ can be updated to approximate z by obtaining a gradient of $\mathcal{F}$ with respect to the belief μ ($\frac{\partial \mathcal{F}}{\partial \mu}$). This update also has to consider the dynamics of the beliefs. The change of the beliefs over time is thus represented by the dynamics and the minimization of the free-energy. This will lead to an approximation of the hidden states $\tilde{z}$ by the beliefs $\tilde{\mu}$ of the system. The update of $\tilde{\mu}$ is mathematically represented as:

$$\dot{\tilde{\mu}} = \frac{d}{dt}\tilde{\mu} - \kappa_\mu \frac{\partial \mathcal{F}}{\partial \tilde{\mu}} \tag{3-13}$$

The update of $\tilde{\mu}$ is based on the gradient descent method often used for minimization problems. The variable κ_μ is the update rate, which in turn is a hyperparameter which can be set to adjust the update of $\tilde{\mu}$.

3-2 Active Inference Control

As stated in the previous section, there are beliefs $\tilde{\mu}$ representing the hidden states $\tilde{z}$ of the system. The dynamics of a belief $\mu^{(i)}$ of the order i is represented by the generative model of the state dynamics $f^{(i)}(\mu^{(i)})$. The basis for determining a control input using Active Inference is by stating that $f^{(i)}(\mu^{(i)})$ is equal to the dynamics needed to reach the desired value.

A common and basic implementation is stating $f(\mu) = z_r - \mu$. This causes the evolution of μ, represented by $f(\mu)$, to converge to a reference value of z_r. The generalized motion represented by μ' and possible higher orders will be updated according to equation (3-13).

When returning to the free-energy equation in equation (3-12), two errors are calculated. First, the error is calculated between the observation o and $g(\mu)$. Second, the error is calculated between the belief μ' and the desired dynamics $f(\mu)$. These two errors are combined in the term $\mathcal{F}$ as shown in equation (3-14). Minimizing $\mathcal{F}$ with respect to μ as in equation (3-13) will thus lead to a value between the hidden state z contained in o and the desired value of the hidden state z contained in $f(\mu)$.

$$\begin{aligned}\mathcal{F} =& \frac{1}{2}\sum_{i=0}^{n_d}\Bigg(\underbrace{(o^{(i)} - g^{(i)}(\mu^{(i)}))^T \Sigma_{o^{(i)}}^{-1}(o^{(i)} - g^{(i)}(\mu^{(i)}))}_{\text{estimation error}} \\ &+ \underbrace{(\mu^{(i+1)} - f^{(i)}(\mu^{(i)}))^T \Sigma_{\mu^{(i)}}^{-1}(\mu^{(i+1)} - f^{(i)}(\mu^{(i)}))}_{\text{desired dynamics error}}\Bigg) + K\end{aligned} \tag{3-14}$$

Note that here, $\Sigma_{\mu^{(i)}}^{-1}$ is now a control hyperparameter instead of the confidence in the process [28]. This is because $f^{(i)}(\mu^{(i)}))$ describes the desired process instead of the true process.

To determine a control input, a different approach to the free-energy equation is taken. The beliefs now have a value which partly contains their desired value. By minimizing $\mathcal{F}$ with respect to o, the observable states can be brought to a value of $g(\mu)$ which partly contains the reference value. When the relation between the observable states o and the control input u is given, $\mathcal{F}$ can be minimized with respect to u.

An update of the control input can be given according to the previous described steps. Note that there are no dynamics of the control input. Thus, the mathematical formulation becomes:

$$\dot{u} = -\kappa_u \frac{\partial \mathcal{F}}{\partial o}\frac{\partial o}{\partial u} \tag{3-15}$$

This introduces an additional hyperparameter κ_u, which influences the update of the control input.

To finalize this in a useable control method for a discrete system, a discretization is required. This discretization is done using a forward Euler method as done in previous research of AIC [1]. For every time instance a new $\tilde{\mu}$ and u is calculated using the known time step T:

$$\begin{aligned} \tilde{\mu}[k+1] &= \tilde{\mu}[k] + \dot{\tilde{\mu}}[k]T \\ u[k+1] &= u[k] + \dot{u}[k]T \end{aligned} \tag{3-16}$$

Note that $\dot{\tilde{\mu}}$ and $\dot{u}$ is calculated according to the values of μ and o at time instance $[k]$.

The AIC control loop can be represented using pseudocode given in Algorithm 1. Throughout this thesis, this algorithm will be referred to as AIC.

Algorithm 1 Active Inference Control

```
Initialize: g^(i)(μ^(i)[0]) = o^(i)[0],    for i = 0, 1, ..., n_d − 1
Input: õ[k]
Output: u[k+1]
μ̃'[k] = d/dt μ̃[k] − κ_μ ∂F/∂μ (μ̃[k], õ[k])        determine dynamics of the belief
μ̃[k+1] = μ̃[k] + μ̃'[k]T                             update beliefs
u'[k] = −κ_u ∂F/∂o ∂o/∂u (μ̃[k], õ[k])               determine dynamics of the control input
u[k+1] = u[k] + u'[k]T                              update control input
save μ̃[k+1], u[k+1]
```

- **Initialize:** $g^{(i)}(\mu^{(i)}[0]) = o^{(i)}[0], \quad \text{for } i = 0, 1, ..., n_d - 1$
- **Input:** $\tilde{o}[k]$
- **Output:** $u[k+1]$
- $\dot{\tilde{\mu}}[k] = \frac{d}{dt}\tilde{\mu}[k] - \kappa_\mu \frac{\partial \mathcal{F}}{\partial \mu}(\tilde{\mu}[k], \tilde{o}[k])$ — determine dynamics of the belief
- $\tilde{\mu}[k+1] = \tilde{\mu}[k] + \dot{\tilde{\mu}}[k]T$ — update beliefs
- $\dot{u}[k] = -\kappa_u \frac{\partial \mathcal{F}}{\partial o}\frac{\partial o}{\partial u}(\tilde{\mu}[k], \tilde{o}[k])$ — determine dynamics of the control input
- $u[k+1] = u[k] + \dot{u}[k]T$ — update control input
- save $\tilde{\mu}[k+1], u[k+1]$

Hyperparameters

The hyperparameters for AIC are the update rates κ_μ and κ_u, and the observation and process noise $\Sigma_o^{(i)}$ and $\Sigma_\mu^{(i)}$. As stated earlier, the parameters $\Sigma_o^{(i)}$ might seem to be based on noise. However, when not known, it can be set such that it dampens stochastic behaviour of the observations [1]. Note that $\Sigma_\mu^{(i)}$ becomes a control hyperparameter with the introduction of desired dynamics.

An intuition of the effect of the hyperparameters can be given when checking the update for $\tilde{\mu}$ in equation (3-13) combined with equation (3-16). The combination in a discrete setting comes down to:

$$\tilde{\mu}[k+1] = \tilde{\mu}[k] + T\left(\frac{d}{dt}\tilde{\mu}[k] - \kappa_\mu \frac{\partial \mathcal{F}[k]}{\partial \tilde{\mu}[k]}\right) \tag{3-17}$$

Assuming $g^{(i)}(\mu^{(i)})$ and $f^{(i)}(\mu^{(i)})$ are linear functions, $\mathcal{F}$ will become a convex quadratic function with respect to μ. The gradient step of $\mathcal{F}$ with respect to μ for the example of a one dimensional μ with only the zeroth order derivative observable then becomes:

$$\frac{\partial \mathcal{F}}{\partial \mu} = -\frac{\partial g}{\partial \mu}\sigma_o^{-1}(o - g(\mu)) + -\frac{\partial f}{\partial \mu}\sigma_\mu^{-1}(\mu' - f(\mu)) \tag{3-18}$$

Note that here σ is used instead of Σ, to represent a one-dimensional scalar instead of a matrix.

Combining the two previous functions allows us to see the effect of the hyperparameters κ_μ, σ_o, and σ_μ.

$$\mu[k+1] = \mu[k] + T\frac{d}{dt}\mu[k] - T\kappa_\mu\left(\frac{\partial g}{\partial \mu}\sigma_o^{-1}(o[k] - g(\mu[k])) + \frac{\partial f}{\partial \mu}\sigma_\mu^{-1}(\mu'[k] - f(\mu[k]))\right) \quad (3\text{-}19)$$

The hyperparameters can be better analysed as the combination $T\kappa_\mu\sigma_o^{-1}$ and $T\kappa_\mu\sigma_\mu^{-1}$. Which when larger than one would update μ with a value larger than the errors, and can cause an overshoot of μ. A bold conclusion can thus be drawn that the combination of $T\kappa_\mu\sigma_o^{-1}$ and $T\kappa_\mu\sigma_\mu^{-1}$ effects the oscillations of the beliefs. Where it seems that $T\kappa_\mu\sigma_o^{-1} < 1$ and $T\kappa_\mu\sigma_\mu^{-1} < 1$ will reduce oscillations of μ, where the lower the value, the more prevention of oscillations. A similar conclusion was also obtained in research done by Baioumy et al. [28]. In addition, they noticed that a larger value would decrease the rise time and thus obtain faster control.

In this thesis, the hyperparameters for which AIC controls a system as fast as possible are not of importance. It is also assumed that the stochastic fluctuations in the observations are unknown. Throughout the simulations and experiment, a standard value of $T = 0.1$ is upheld. Therefore, to reduce oscillations, values of $\kappa_\mu = 1$ and $\kappa_u = 1$ are set. For the values of $\Sigma_{\mu^{(i)}}$ and $\Sigma_{o^{(i)}}$, an identity matrix is set with the dimension equal to that of $\mu^{(i)}$ and $o^{(i)}$. This causes $T\kappa_\mu\sigma_o^{-1} = 0.1$ and $T\kappa_\mu\sigma_\mu^{-1} = 0.1$ and will drastically prevent oscillations. It also indicates that there is no favourable bias to either update μ to be a better estimator or cause faster control.

For tuning of these hyperparameters, I refer to previous research done by Pezzato et al. and Baioumy et al. [1, 28].

3-3 Active Inference Control For Nonholonomic System

To apply AIC to a differential-drive two-wheeled mobile robot, first the known observations need to be stated and their relationship to the hidden states. The 2D workspace was defined with the three states $z = [x\ y\ \theta]^T$ in chapter 2. As stated in assumption 2, it is assumed that these three states and therefore their velocities $\dot{z}$ since u is also known. Therefore, $n_d = 1$ and the observations are defined as $\tilde{o} = [o_x\ o_y\ o_\theta\ o'_x\ o'_y\ o'_\theta]^T$. It is assumed the observed data contains Gaussian noise in addition to the true state value, but the value of the variance of this noise is not known.

The three states x, y, and θ and their velocities are also the only states needed for the representation of the dynamics, therefore these also represent all the hidden states $\tilde{z} = [z^T\ \dot{z}^T]^T$. The representation of $\tilde{z}$ is used for the true states. The observations' and hidden states' relationship is defined, and therefore the following can be stated:

$$\underbrace{\tilde{z}}_{g(\tilde{z})} + \begin{bmatrix} w_o \\ w_o' \end{bmatrix} = \begin{bmatrix} x(t) \\ y(t) \\ \theta(t) \\ cos(\theta(t))v(t) \\ sin(\theta(t))v(t) \\ \omega(t) \end{bmatrix} + \begin{bmatrix} w_o \\ w_o' \end{bmatrix} = \tilde{o} = \begin{bmatrix} o_x(t) \\ o_y(t) \\ o_\theta(t) \\ cos(o_\theta(t))v(t) \\ sin(o_\theta(t))v(t) \\ \omega(t) \end{bmatrix}, \qquad u = \begin{bmatrix} v(t) \\ \omega(t) \end{bmatrix} \tag{3-20}$$

As stated in the previous section, a belief $\tilde{\mu}$ can be used to approximate the true states $\tilde{z}$. Therefore, the relation between $\tilde{z}$ and $\tilde{o}$ is the same as the relation between $\tilde{\mu}$ and $\tilde{o}$. The goal is to control the system to the origin of the three-dimensional space of z. For that reason, the dynamics of μ are set as:

$$\begin{aligned} f(\mu) &= 0 - \mu \\ f'(\mu') &= -\mu' \end{aligned} \tag{3-21}$$

The observations, their relationship to the beliefs, and the dynamics of the beliefs, are all defined. This allows us to solve the variational free-energy defined in (3-12), the K term cannot be minimized and therefore not fully shown.

$$\begin{aligned} \mathcal{F} =& \frac{1}{2}\sum_{i=0}^{1}\Big((o^{(i)} - g^{(i)}(\mu^{(i)}))^T\Sigma_{o^{(i)}}^{-1}(o^{(i)} - g^{(i)}(\mu^{(i)})) \\ &+(\mu^{(i+1)} - f^{(i)}(\mu^{(i)}))^T\Sigma_{\mu^{(i)}}^{-1}(\mu^{(i+1)} - f^{(i)}(\mu^{(i)}))\Big) + K \end{aligned} \tag{3-22}$$

The error between the observations and the beliefs, and between the beliefs and their desired dynamics are defined as:

$$\begin{aligned} o - g(\mu) &= \begin{bmatrix} o_x \\ o_y \\ o_\theta \end{bmatrix} - \begin{bmatrix} \mu_x \\ \mu_y \\ \mu_\theta \end{bmatrix}, & o' - g'(\mu') &= \begin{bmatrix} cos(o_\theta)v \\ sin(o_\theta)v \\ \omega \end{bmatrix} - \begin{bmatrix} \mu_x' \\ \mu_y' \\ \mu_\theta' \end{bmatrix} \\ \mu' - f(\mu) &= \begin{bmatrix} \mu_x' \\ \mu_y' \\ \mu_\theta' \end{bmatrix} - \begin{bmatrix} 0-\mu_x \\ 0-\mu_y \\ 0-\mu_\theta \end{bmatrix}, & \mu'' - f'(\mu') &= \begin{bmatrix} \mu_x'' \\ \mu_y'' \\ \mu_\theta'' \end{bmatrix} - \begin{bmatrix} -\mu_x' \\ -\mu_y' \\ -\mu_\theta' \end{bmatrix} \end{aligned} \tag{3-23}$$

The update rule $\dot{\tilde{\mu}}$ can be defined since the effect of $\tilde{\mu}$ is now known for $\mathcal{F}$.

$$\begin{aligned} \dot{\tilde{\mu}} &= \begin{bmatrix} \dot{\mu} \\ \dot{\mu}' \\ \dot{\mu}'' \end{bmatrix} = \frac{d}{dt}\mu - \kappa_\mu \frac{\partial \mathcal{F}}{\partial \mu} \\ \dot{\tilde{\mu}} &= \begin{bmatrix} \mu' \\ \mu'' \\ 0 \end{bmatrix} - \kappa_\mu\left(\begin{bmatrix} \Sigma_o^{-1} & 0 \\ 0 & \Sigma_{o'}^{-1} \\ 0 & 0 \end{bmatrix}\begin{bmatrix} o-\mu \\ o'-\mu' \end{bmatrix} + \begin{bmatrix} \Sigma_\mu^{-1} & 0 \\ \Sigma_\mu^{-1} & \Sigma_{\mu'}^{-1} \\ 0 & \Sigma_{\mu'}^{-1} \end{bmatrix}\begin{bmatrix} \mu' - (0-\mu) \\ \mu'' + \mu' \end{bmatrix}\right) \\ &= \begin{bmatrix} \mu' \\ \mu'' \\ 0 \end{bmatrix} - \left(\begin{bmatrix} o-\mu \\ o'-\mu' \\ 0 \end{bmatrix} + \begin{bmatrix} \mu' - (0-\mu) \\ (\mu''+\mu') + (\mu' - (0-\mu)) \\ \mu'' + \mu' \end{bmatrix}\right) \end{aligned} \tag{3-24}$$

To give some general intuition of how the belief will develop, the hyperparameters are all set to one or the identity matrix as stated in section 3-2. What can be inferred from this equation is that μ will be in between the observations o and the desired value zero. The same is for μ' which will be updated to be in between the values of o' and $-\mu$. μ'' is a variable which will adapt to be the negative value of μ' and dampen the update of μ'.

The relationship of the observations and the control input needs to be known to be able to determine the control input as required in equation (3-26). At the start of this section $\tilde{o}$ was defined and control input u (equation (3-20)). This allows us to obtain:

$$\frac{\partial \tilde{o}}{\partial u} = \begin{bmatrix} \frac{\partial \tilde{o}}{\partial v} & \frac{\partial \tilde{o}}{\partial \omega} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ cos(o_\theta) & 0 \\ sin(o_\theta) & 0 \\ 0 & 1 \end{bmatrix} \tag{3-25}$$

It is also important to give some general intuition for the update of the control input. Again, all Σ matrices are set to the identity matrix, and are therefore omitted from the next equation. Using the given equations, the control input is updated according to:

$$\begin{aligned} \dot{u} &= -\kappa_u \begin{bmatrix} cos(o_\theta)(cos(o_\theta)v - \mu'_x) + sin(o_\theta)(sin(o_\theta)v - \mu'_y) \\ \omega - \mu'_\theta \end{bmatrix} = \begin{bmatrix} \dot{v} \\ \dot{\omega} \end{bmatrix} \\ \begin{bmatrix} v[k+1] \\ \omega[k+1] \end{bmatrix} &= \begin{bmatrix} v[k] \\ \omega[k] \end{bmatrix} + \begin{bmatrix} \dot{v}[k] \\ \dot{\omega}[k] \end{bmatrix} T \end{aligned} \tag{3-26}$$

Although not immediately apparent from the given update equations, it is impossible to converge a two-wheeled differential drive mobile robot to a three-dimensional reference state from <u>every</u> initial state. Allow me to elaborate on this with one example of an initial state for which it will not work.

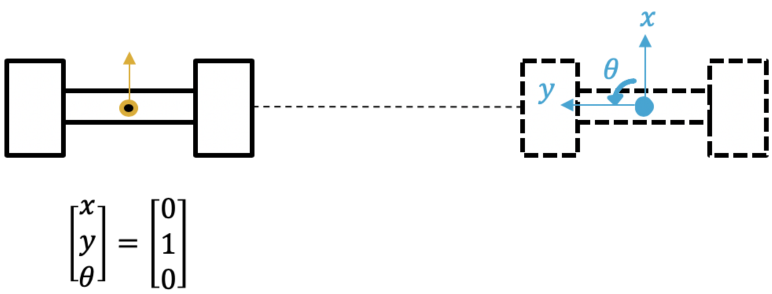


Figure 3-1: The two-wheeled robot on the left is the starting position, the dashed robot on the right is the desired position also containing the origin.

Assume the starting position of the robot according to Figure 3-1. To make this example simple, an assumption is made that we have the exact values of $\tilde{z}$. Four of the six states have already reached their desired value. The belief $\tilde{\mu}$ has been initiated with correct values for $\mu_x, \mu_\theta, \mu'_x, \mu'_\theta$ and has stable values for μ_y, μ'_y.

$$\tilde{z} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \tilde{\mu} = \begin{bmatrix} 0 \\ \frac{2}{3} \\ 0 \\ 0 \\ \frac{1}{3} \\ 0 \end{bmatrix} \tag{3-27}$$

It is clear that only the y state needs to be controlled to a value of zero. The values can be used for the equation of the update rule of the control input. Filling in the variables in equation (3-26) will give the following:

$$\dot{u} = -\kappa_u \begin{bmatrix} 1 \cdot (1 \cdot 0 - 0) + 0 \cdot (0 \cdot 0 - \frac{1}{3}) \\ 0 - 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{3-28}$$

The control input will not increase nor decrease. When a control input of 0 is initialized, the robot will not move to the origin. The issue is the constraint applicable for this nonholonomic system. The velocity in y direction is zero when $\theta = 0$ because the system upholds to $\dot{x}(t)sin(\theta(t)) - \dot{y}(t)cos(\theta) = 0$.

The issue is caused by the controller controlling the value of θ to zero and keeping it there. The rate at which θ goes to zero is determined by (3-29).

$$\begin{aligned} \theta(t) &= \theta(0) + \int_0^t \omega(t)\ dt \\ \omega(t) &= \omega(0) + \int_0^t \dot{\omega}(t)\ dt \\ &= \omega(0) - \kappa_u \int_0^t \omega(t) - \mu'_\theta(t)\ dt \end{aligned} \tag{3-29}$$

Note that μ'_θ is only based on the observation of θ when looking at previous equations, and knowing that μ_θ is only based on θ according to equations (3-23) and (3-24). The control of θ thus does not take the values of x or y into account.

However, the control of y is based on the value of $\theta(t)$ and $y(t)$ as:

$$\begin{aligned} y(t) &= y(0) + \int_0^t sin(o_\theta(t))v(t)\ dt \\ v(t) &= v(0) + \int_0^t \dot{v}(t)\ dt \\ &= v(0) - \kappa_u \int_0^t cos(o_\theta(t))(cos(o_\theta(t))v(t) - \mu'_x(t)) + sin(o_\theta(t))(sin(o_\theta(t))v(t) - \mu'_y(t))\ dt \end{aligned} \tag{3-30}$$

What happens is that θ influences the control of y, while y does not influence the control of θ.

In order for the control method to converge to the origin of all three dimensions, it is required that the control of θ is influenced by both x and y. Adjusting this also corresponds to the connection between three states stated in the nonholonomic constraint $\dot{x}(t)sin(\theta(t)) - \dot{y}(t)cos(\theta(t)) = 0$.

In the next section, Hierarchical Active Inference Control (HAIC) is explained which adds this connection of x and y to the control of θ.

3-4 Hierarchical Active Inference Control

In the previous section, it was shown that AIC was unable to control a nonholonomic system to a three-dimensional reference point. The main cause of the inability was the constraint of the nonholonomic system causing the evolution of the three states to be dependent on one another. Failure was caused by AIC not taking this dependency into account. In this section, a solution making use of Hierarchical Active Inference Control (HAIC) is proposed which can incorporate the constraint such that the controller can control the robot to the origin.

The Nonholonomic Constraint

First let us rewrite the nonholonomic constraint of a two-wheeled differential-drive robot as follows:

$$\frac{\dot{y}(t)}{\dot{x}(t)} = \frac{sin(\theta(t))}{cos(\theta(t))} = tan(\theta(t)) \tag{3-31}$$

This nonholonomic constraint can be related to the belief μ. According to the previous section, the beliefs μ'_x and μ'_y are partly updated to represent the dynamics necessary to reach the origin. Therefore, these beliefs can be used to obtain an additional belief. This belief represents the orientation best suitable for the dynamics in the x and y direction. The belief of this orientation will be indicated using $\mu_{r,\theta}$. The following equation represents the relation between the beliefs:

$$\mu_{r,\theta} = arctan\left(\frac{\mu'_y}{\mu'_x}\right) \tag{3-32}$$

Making It Hierarchical

As done in other research, the Active Inference framework can be used to give layered structure to control [31, 30]. When doing this with AIC an AIC layer only makes use of the beliefs of the layer below and regulates the lower layer by giving them a reference value (indicated with μ_r).

One can implement an additional layer of AIC on top of the current one as shown in Figure 3-2. The top AIC layer uses the beliefs of the bottom layer as observations (o_T), and has its own generative models (g_T, f_T) as explained in section 3-1. The subscript T indicates the generative model belongs to the top AIC layer.

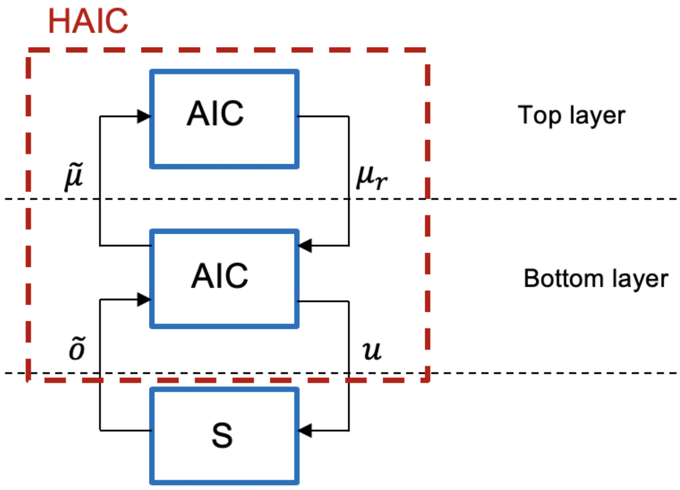


Figure 3-2: Block Scheme representing HAIC. The bottom layer of AIC uses observations ($\tilde{o}$) and gives a control input (u). The top layer makes use of the beliefs of the bottom layer ($\tilde{\mu}$), and gives a reference belief (μ_r) to the bottom layer.

The generative models for the top layer are defined as:

$$\begin{aligned} g_T(\mu_{r,\theta}) + w_{o^T} = \mu_r + w_{o^T} = o_T = arctan\left(\frac{\mu'_y}{\mu'_x}\right), \qquad & w_{o^T} \sim (0, \sigma_{o^T}) \\ f_T(\mu_{r,\theta}) + w_{\mu_{r,\theta}} = 0 - \mu_{r,\theta} + w_{\mu_{r,\theta}} = \mu'_{r,\theta}, \qquad & w_{\mu_{r,\theta}} \sim (0, \sigma_{\mu_{r,\theta}}) \end{aligned} \tag{3-33}$$

Because o_T and $\mu_{r,\theta}$ are one dimensional, σ is used as a representation of the variance. Note that there is no derivative available of the angle. Therefore, the generalized motions of the top layer are up to an order of $n_d = 0$.

The variances $\sigma_{o_T}^{-1}$ and $\sigma_{\mu_r}^{-1}$ cannot be related to the noise similarly as those of the bottom layer AIC. To elaborate on what the confidences can represent, let us first look at how they will influence the belief by determining the free energy equation for the top layer. From the observation and the general motions, the free energy equation is given as:

$$\mathcal{F}_T = \frac{1}{2}\left(\left(arctan\left(\frac{\mu'_y}{\mu'_x}\right) - \mu_{r,\theta}\right)\sigma_{o_T}^{-1}\left(arctan\left(\frac{\mu'_y}{\mu'_x}\right) - \mu_{r,\theta}\right) + \left(\mu_{r,\theta} + \mu'_{r,\theta}\right)\sigma_{\mu_{r,\theta}}^{-1}\left(\mu_{r,\theta} + \mu'_{r,\theta}\right)\right) + K \tag{3-34}$$

Updating the belief according to equation (3-13) will give:

$$\begin{bmatrix} \dot{\mu}_{r,\theta} \\ \dot{\mu}'_{r,\theta} \end{bmatrix} = \begin{bmatrix} \mu'_{r,\theta} \\ 0 \end{bmatrix} - \kappa_{\mu_r}\begin{bmatrix} -\sigma_{o_T}^{-1}\left(arctan\left(\frac{\mu'_y}{\mu'_x}\right) - \mu_{r,\theta}\right) + \sigma_{\mu_{r,\theta}}^{-1}(\mu_{r,\theta} + \mu'_{r,\theta}) \\ \sigma_{\mu_{r,\theta}}^{-1}(\mu_{r,\theta} + \mu'_{r,\theta}) \end{bmatrix} \tag{3-35}$$

The update of $\tilde{\mu}_{r,\theta} = [\mu_{r,\theta}\ \mu'_{r,\theta}]^T$ would lead to a value between o_T and 0 for $\mu_{r,\theta}$. The system will neither reach the origin in the x and y direction nor in the orientation of zero when o_T is non-zero. Ideally, the robot would reach the origin in the x and y direction, and after it will turn its orientation to zero. To obtain this, $\sigma_{o_T}^{-1}$ and $\sigma_{\mu_{r,\theta}}^{-1}$ can be used to change when the reference orientation focuses on minimizing the x and y distance or the reference focuses on bringing θ to zero.

The effect of $\sigma_{o_T}^{-1}$ is that it will update $\mu_{r,\theta}$ to $arctan(\mu_y'/\mu_x')$. Pointing the robot in the direction such that it can reduce the error in x and y dimension. The effect of $\sigma_{\mu_{r,\theta}}^{-1}$ is that it will update $\mu_{r,\theta}$ to zero. A solution is stating that $\sigma_{o_T}^{-1} = 1$ when the origin in the x and y plane has not been reached, and $\sigma_{\mu_{r,\theta}}^{-1} = 0$. And stating that $\sigma_{\mu_{r,\theta}}^{-1} = 1$ when it has been reached, and $\sigma_{\mu_{r,\theta}}^{-1} = 0$.

The values of σ_{o_T} and $\sigma_{\mu_{r,\theta}}$ relate to confidence of being able to converge to the desired set point. Where we can say that $\sigma_{o_T} = 1$ when the controller is confident that it has to reduce the error in the x and y dimension, and $\sigma_{o_T} = \infty$ when it does not have to. Vice versa is applicable for $\sigma_{\mu_{r,\theta}}$, where $\sigma_{\mu_{r,\theta}} = \infty$ when reducing the error of x and y and $\sigma_{\mu_{r,\theta}} = 1$ when converging θ to zero.

Introduction Of The Decision Parameter

The best way is to have a binary switch for the transition of the values $\sigma_{o_T}^{-1}$ and $\sigma_{\mu_{r,\theta}}^{-1}$ for usage in equation (3-35). This switch is based on the distance in the x and y direction contained in the beliefs μ_x' and μ_y'. This allows for a direct change by having the values of $\sigma_{o_T}^{-1}$ and $\sigma_{\mu_{r,\theta}}^{-1}$ immediately switch to one or zero.

This binary step for $\sigma_{o_T}^{-1}$ and $\sigma_{\mu_{r,\theta}}^{-1}$ is defined as:

$$
\begin{aligned}
\sigma_{o_T}^{-1} &= \begin{cases} 0 & \text{if } \sqrt{{\mu_y'}^2 + {\mu_x'}^2} < \epsilon \\ 1 & \text{if } \sqrt{{\mu_y'}^2 + {\mu_x'}^2} \geq \epsilon \end{cases} \\
\sigma_{\mu_{r,\theta}}^{-1} &= \begin{cases} 1 & \text{if } \sqrt{{\mu_y'}^2 + {\mu_x'}^2} < \epsilon \\ 0 & \text{if } \sqrt{{\mu_y'}^2 + {\mu_x'}^2} \geq \epsilon \end{cases}
\end{aligned} \tag{3-36}
$$

Here ϵ is a decision variable indicating when the values for $\sigma_{o_T}^{-1}$ and $\sigma_{\mu_{r,\theta}}^{-1}$ will switch. Note that zero is an approximation of ∞^{-1}.

Other transitions that are piecewise continuous can give settling points between the values of o_T and zero. Any other piecewise continuous switch between zero and one for $\sigma_{o_T}^{-1}$ and $\sigma_{\mu_{r,\theta}}^{-1}$ can give an update to a random orientation. This does not reduce the error in x and y direction but also not bring $\mu_{r,\theta}$ to a value of zero. This can be seen by rewriting the update equation to:

$$
\begin{aligned}
\frac{\partial \mathcal{F}_T}{\partial \mu_{r,\theta}} =& -\sigma_{o_T}^{-1}\left(arctan\left(\frac{\mu_y'}{\mu_x'}\right) - \mu_{r,\theta}\right) + \sigma_{\mu_{r,\theta}}^{-1}(\mu_{r,\theta} + \mu_{r,\theta}') \\
=& \mu_{r,\theta}\underbrace{(\sigma_{o_T}^{-1} + \sigma_{\mu_{r,\theta}}^{-1})}_{=1} - \underbrace{\left(\sigma_{o_T}^{-1} arctan\left(\frac{\mu_y'}{\mu_x'}\right) + \sigma_{\mu_{r,\theta}}^{-1}\mu_{r,\theta}'\right)}_{\text{becomes a random orientation}}
\end{aligned} \tag{3-37}
$$

The selection of ϵ is based on the values of $\tilde{\mu}$ when the system is at the origin. Obtaining a value of exactly zero for any value of $\tilde{\mu}$ is highly unlikely. Even when the observations o

become exactly zero, $\tilde{\mu}$ will still be updated to become close to zero but not exactly zero. Though as stated earlier, it is assumed the observations contain some Gaussian noise. For that reason, it is required to set ϵ to some positive digit such that it considers the noise transferred from the observations to $\tilde{\mu}$. Where the higher the digit, the less sensitive the switching of $\sigma_{o_T}^{-1}$ and $\sigma_{\mu_{r,\theta}}^{-1}$ will be to noise contained in $\tilde{\mu}$.

On the other hand, setting ϵ to a high value will cause the HAIC to control the system to an angle of zero for high values of $\tilde{\mu}$. When θ is controlled to zero, no evolution in the y dimension possible. This can be seen when applying $\theta = 0$ to the nonholonomic constraint: $cos(\theta)\dot{y} - sin(\theta)\dot{x} = 0$. Here, $\dot{y}$ is zero when $\theta = 0$. Therefore, the higher ϵ the higher a possible error in the y dimension.

The selection of ϵ is a trade-off between robustness against noise and a possible error in the y dimension. In chapter 4, the effect of the hyperparameter ϵ is shown for different level of noise present in the measurement. First, the error is compared for two different values of ϵ. Second, noise in the observations is added such that the switching of $\sigma_{o_T}^{-1}$ and $\sigma_{\mu_{r,\theta}}^{-1}$ can be checked. Based on these simulations, an intuition about the correct selection of the parameter is given.

Bringing It Together

To bring everything together, the effect of the top layer on the bottom layer is described. The top AIC layer gives a reference to the bottom layer AIC. The bottom layer AIC is similar to the one described in section 3-3 except for the generative model of the state dynamic of μ_θ which is:

$$f(\mu_\theta) = \mu_{r,\theta} - \mu_\theta \tag{3-38}$$

The following update rules for $\tilde{\mu}$, $\tilde{\mu}_{r,\theta}$ and u are then given in equation (3-39). The evolution of all beliefs and the control input are written as:

$$\begin{aligned}
\dot{\tilde{\mu}} &= \begin{bmatrix} \mu' \\ \mu'' \\ 0 \end{bmatrix} + \kappa_\mu \left(\begin{bmatrix} \Sigma_o^{-1} & 0 \\ 0 & \Sigma_{o'}^{-1} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} o - \mu \\ o' - \mu' \end{bmatrix} + \begin{bmatrix} \Sigma_\mu^{-1} & 0 \\ \Sigma_\mu^{-1} & \Sigma_{\mu'}^{-1} \\ 0 & \Sigma_{\mu'}^{-1} \end{bmatrix} \begin{bmatrix} \mu' + \mu - \begin{bmatrix} 0 & 0 & \mu_{r,\theta} \end{bmatrix}^T \\ \mu'' + \mu' \end{bmatrix} \right) \\
\dot{\tilde{\mu}}_{r,\theta} &= \begin{bmatrix} \mu'_{r,\theta} \\ 0 \end{bmatrix} - \kappa_{\mu_r} \begin{bmatrix} -\sigma_{o_T}^{-1} \left(arctan\left(\frac{\mu'_y}{\mu'_x}\right) - \mu_{r,\theta} \right) + \sigma_{\mu_{r,\theta}}^{-1}(\mu_{r,\theta} + \mu'_{r,\theta}) \\ \sigma_{\mu_{r,\theta}}^{-1}(\mu_{r,\theta} + \mu'_{r,\theta}) \end{bmatrix} \\
\dot{u} &= -\kappa_u \begin{bmatrix} cos(\theta(t))\sigma_{o'_x}(cos(\theta(t))v(t) - \mu'_x) + sin(\theta(t))\sigma_{o'_y}(sin(\theta(t))v(t) - \mu'_y) \\ \sigma_{o'_\theta}(\omega(t) - \mu'_\theta) \end{bmatrix}
\end{aligned} \tag{3-39}$$

To summarize HAIC and apply it to a nonholonomic system, a description in algorithm form is given. For the use on a discrete system, the following algorithm in pseudocode is given in algorithm 2. This algorithm is implemented throughout this thesis.

Algorithm 2 Hierarchical Active Inference Control

Initialize: $g^{(i)}(\mu^{(i)}[0]) = o^{(i)}[0], \quad \text{for } i = 0, 1, ..., n_d - 1$
Input: $\tilde{o}[k]$
Output: $u[k+1]$
$\dot{\tilde{\mu}}[k] = \frac{d}{dt}\tilde{\mu}[k] - \kappa_\mu \frac{\partial \mathcal{F}}{\partial \mu}(\tilde{\mu}[k], \tilde{o}[k], \mu_{r,\theta}[k])$ determine dynamics of bottom layer beliefs
$\tilde{\mu}[k+1] = \tilde{\mu}[k] + \dot{\tilde{\mu}}[k]T$ update the bottom layer beliefs
$o_T = arctan\left(\frac{\mu'_y}{\mu'_x + 10^{-5}}\right)$ $10^{-}5$ is used to not divide by zero
$\dot{\tilde{\mu}}_{r,\theta}[k] = \frac{d}{dt}\tilde{\mu}_{r,\theta}[k] - \kappa_{\mu_{r,\theta}} \frac{\partial \mathcal{F}_T}{\partial \mu}(\tilde{\mu}_{r,\theta}[k], \tilde{\mu}[k])$ determine dynamics of the top layer beliefs
$\tilde{\mu}_{r,\theta}[k+1] = \tilde{\mu}_{r,\theta}[k] + \dot{\tilde{\mu}}_{r,\theta}[k]T$ update the top layer beliefs
$\dot{u}[k] = -\kappa_u \frac{\partial \mathcal{F}}{\partial o}\frac{\partial o}{\partial u}(\tilde{\mu}[k], \tilde{o}[k])$ determine dynamics of the control input
$u[k+1] = u[k] + \dot{u}[k]T$ update the control input
save $\tilde{\mu}[k+1], \tilde{\mu}_{r,\theta}[k+1], u[k+1]$

3-5 Conclusion

In this chapter, the framework of Active Inference was explained and the steps to obtaining a control input. The four hyperparameters generally present in AIC are set to one or the identity matrix throughout this Thesis. This prevents oscillations of the states when using time step of $0.1s$, as done throughout the simulations and experiment described in this report.

It has been shown that AIC is unable to control a two-wheeled derivative mobile robot to a three-dimensional reference state. This was because applying AIC directly to the system will not take the relation between x, y, and θ present for nonholonomic systems into account. Controlling θ to the desired value while the control of x and y are also based on the value for θ. Thus, answering the first sub research question of "What are the limitations of AIC for controlling a nonholonomic mobile robot to a three-dimensional reference state?"

A solution to the problem is the introduction of another layer of AIC, thus creating HAIC. This additional layer adds a relation between x, y, and θ based on the nonholonomic constraint of the system. It also introduces an additional hyperparameter ϵ. The hyperparameter ϵ determines when θ is controlled such that the error in the x and y direction is reduced, or when it is controlled to a value of zero.

The selection of the hyperparameter ϵ comes with a trade-off. Either it can be chosen to allow a larger error in the y dimension and have more robust control against noise present in the observations. Or a decision can be made to obtain a minimal error in the y dimension but have the possibility to not be able to converge to the origin when too much noise is present. This trade-off is discussed in the next chapter where the effect of ϵ is discussed.

Chapter 4

Simulations

In this chapter, two simulations are used to verify whether Hierarchical Active Inference Control (HAIC) can converge a two-wheeled differential-drive mobile robot to a three-dimensional reference point.

The first simulations are run with varying starting points. The goal is to check whether it can always converge and to check how the controller responds to varying starting points. These simulations do not contain any additive noise and thus assume perfect conditions, measurements, and system dynamics.

The second simulation adds noise to the observations used in HAIC. The reason for this addition is to confirm whether noise affects the selection of the decision variable ϵ which determines the values of $\sigma_{o_T}^{-1}$ and $\sigma_{\mu_{r,\theta}}^{-1}$. As stated in the previous chapter, a higher value for ϵ will make the controller more robust against noise. The simulation is used to verify whether HAIC can converge when the value of ϵ is increased.

As stated in the previous chapter, the following hyperparameters for Active Inference Control (AIC) and HAIC are set to values of 1 or the identity matrix I:

$$
\begin{aligned}
&\kappa_u = \kappa_\mu = \kappa_{\mu_{r,\theta}} = 1 \\
&\Sigma_o = \Sigma_{o'} = \Sigma_\mu = \Sigma_{\mu'} = \Sigma_{\mu''} = I \in \mathbb{R}^{3\times 3}
\end{aligned}
\tag{4-1}
$$

All simulations are run with a time step of $T = 0.1$ s. When referring to a time stamp such as $t = 10$ s it is always referred to the time instance $k = 10t$.

4-1 Varying Starting Positions

For the first simulation, varying starting points are simulated for the controlled system. The variations are based on a variable φ given in degrees($^\circ$). This variable determines the starting position based on:

$$z[0] = \begin{bmatrix} cos(\varphi) \\ sin(\varphi) \\ \varphi \end{bmatrix} \tag{4-2}$$

The simulations are equally spread over $0 \leq \varphi \leq 180^\circ$ (1801 in total), thus obtaining data for $\varphi = 0.0^\circ, 0.1^\circ, 0.2^\circ, ..., 180.0^\circ$. A figure showing the half circle of starting positions is given in Figure 4-1.

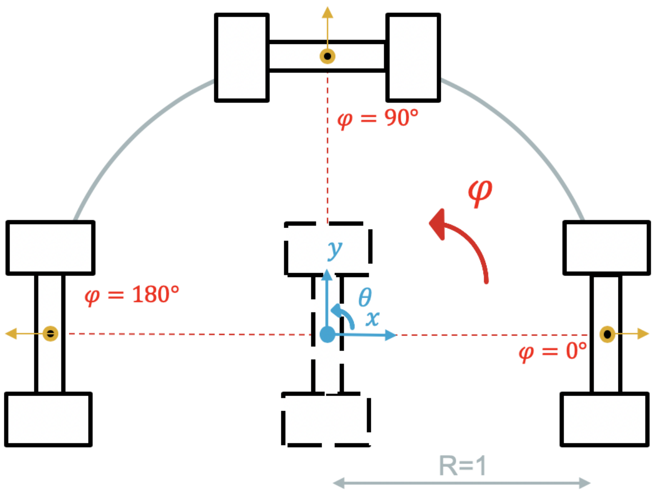


Figure 4-1: The variation of φ and the effect it has on the starting position of the mobile robot. The dashed robot with the blue arrows is the origin of the three-dimensional space and also the goal position for every variation of φ. The grey line indicates the line on which the starting positions are positioned, it is a half circle with a radius of $1\ m$. The yellow arrow of the full robot indicates the starting orientation of the robot, it always points outwards with respect to the half circle. The three robots with the yellow arrows indicate the starting positions of $\varphi = 180^\circ$, $\varphi = 90^\circ$, and $\varphi = 0$

To summarize the response of the controllers for the simulations, two performance metrics are used. The first performance metric is the cost of the system. The cost at a certain step is indicated as $c[k]$, or $c(t)$ when using a time stamp. The cumulative cost J of the control of the system from step 0 until N is defined as in (4-3). N is the number of iterations it takes for the system to converge. The system is assumed to be converged when the values for $|x|$, $|\dot{x}|$, $|\theta|$, $|\dot{\theta}|$ are all lower than 0.001. The dimension y is omitted since the convergence can gain a small error as stated in the previous chapter. The second performance metric is the time it requires to converge to a stable position close to the origin.

$$J = \sum_{k=0}^{N} c[k] = \sum_{k=0}^{N} z^T[k]Qz[k] + u^T[k]Ru[k]$$
$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\pi^2} \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4-3}$$

The value of $1/\pi^2$ is used in (4-3) to normalize the error for all dimensions. The max starting error of the x, y and θ dimension, as part of the cost function $c[k]$, are now all set to one.

Simulations with HAIC for two different values of ϵ and Model Predictive Control (MPC) (according to [23] with optimal hyperparameters set based on the same paper) are run for all values of φ. They are applied to a two-wheeled differential-drive mobile robot with the dynamics stated in chapter 2 according to (2-2).

In Figure 4-2, the results are shown against the values of φ. The value J increases with a higher value for φ, which can be expected since a higher value for φ also increases the starting value of $c[o]$. When we compare the values for HAIC and MPC it can be seen that HAIC clearly underperforms compared to MPC. Both the cost and the convergence time are higher for HAIC. Although the goal was never to have HAIC outperform MPC, it can be concluded HAIC does not obtain an optimal trajectory or control input.

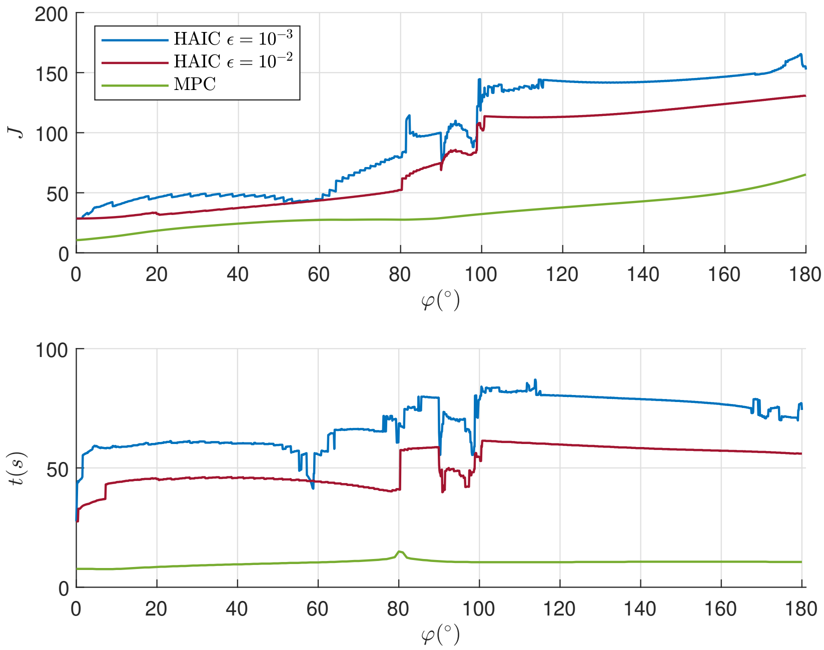


Figure 4-2: Graphs showing the performance of HAIC and MPC where (TOP) the value of the cost function J and (BOTTOM) the convergence time. The system is assumed to be converged when the values for $|x|$, $|\dot{x}|$, $|\theta|$, $|\dot{\theta}|$ are all lower than 0.001

When comparing the two HAIC for different values of ϵ, it can be seen that a higher value for ϵ causes a smoother value for J for small changes in φ. Though a large change in J

can be seen for values of $80^\circ \leq \varphi \leq 105^\circ$ for both values of ϵ. This is likely caused by the system operating at the border of the function of $arctan(y/x)$ which is used in the top layer of HAIC. Where the value of $arctan(y/x)$ can shift drastically when a small positive value for x changes to a small negative value. Causing y/x to go from ∞ to $-\infty$ when y stays negative or positive.

A convergence time for a higher value of ϵ is to be expected because the HAIC decides to control the system to $\theta = 0$ while accepting a larger error in the y dimension. When the convergence time can be checked against the final error ($c[N]$) in Figure 4-3. It can be seen that a tradeoff between convergence time and final error is made when choosing a value for ϵ. The higher ϵ, the lower the convergence time but the higher the final error.

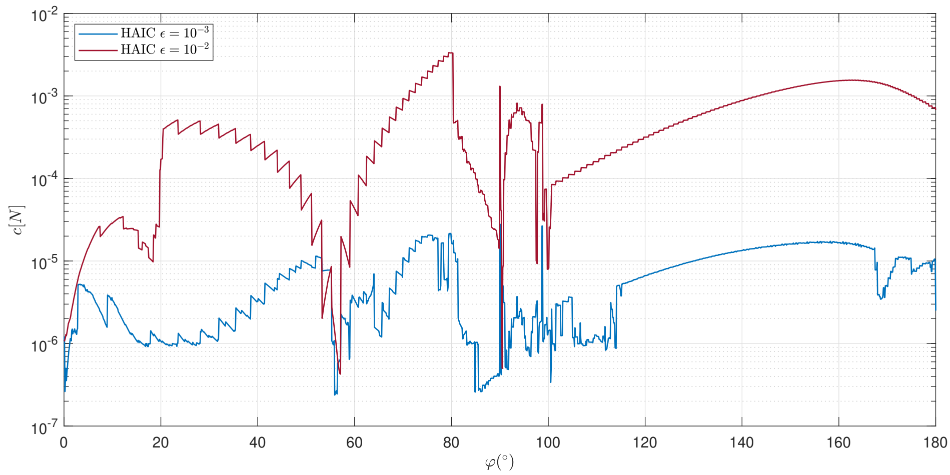


Figure 4-3: Graphs showing the final value $c[n]$ of two HAIC controllers for values of $\epsilon = 10^{-2}$ and $\epsilon = 10^{-3}$. Note that the y-axis is on a logarithmic scale. When compared to Figure 4-2, this graph contains the cost at the last time stamp when the system is converged, instead of the summation of the cost over the full convergence time.

In Figure 4-4 the cost J for HAIC with a value of $\epsilon = 10^{-3}$ as stated in equation (4-3) is split for the x and y term, the θ term, and the u term of the function. The reason an analysis is done for this value of ϵ is because of the higher changes of J for different values φ. The splits are represented as given in equation (4-4). The reason for these splits is such that an identification can be done of what is causing the large variation in cost for small variations of φ.

$$\begin{aligned}
J &= J_{x,y} + J_\theta + J_u \\
J_{x,y} &= \sum_{k=0}^{N} \begin{bmatrix} x[k] & y[k] \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x[k] \\ y[k] \end{bmatrix} \\
J_\theta &= \frac{1}{\pi^2} \sum_{k=0}^{N} (\theta[k])^2 \\
J_u &= \sum_{k=0}^{N} \begin{bmatrix} v[k] & \omega[k] \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v[k] \\ \omega[k] \end{bmatrix}
\end{aligned} \tag{4-4}$$

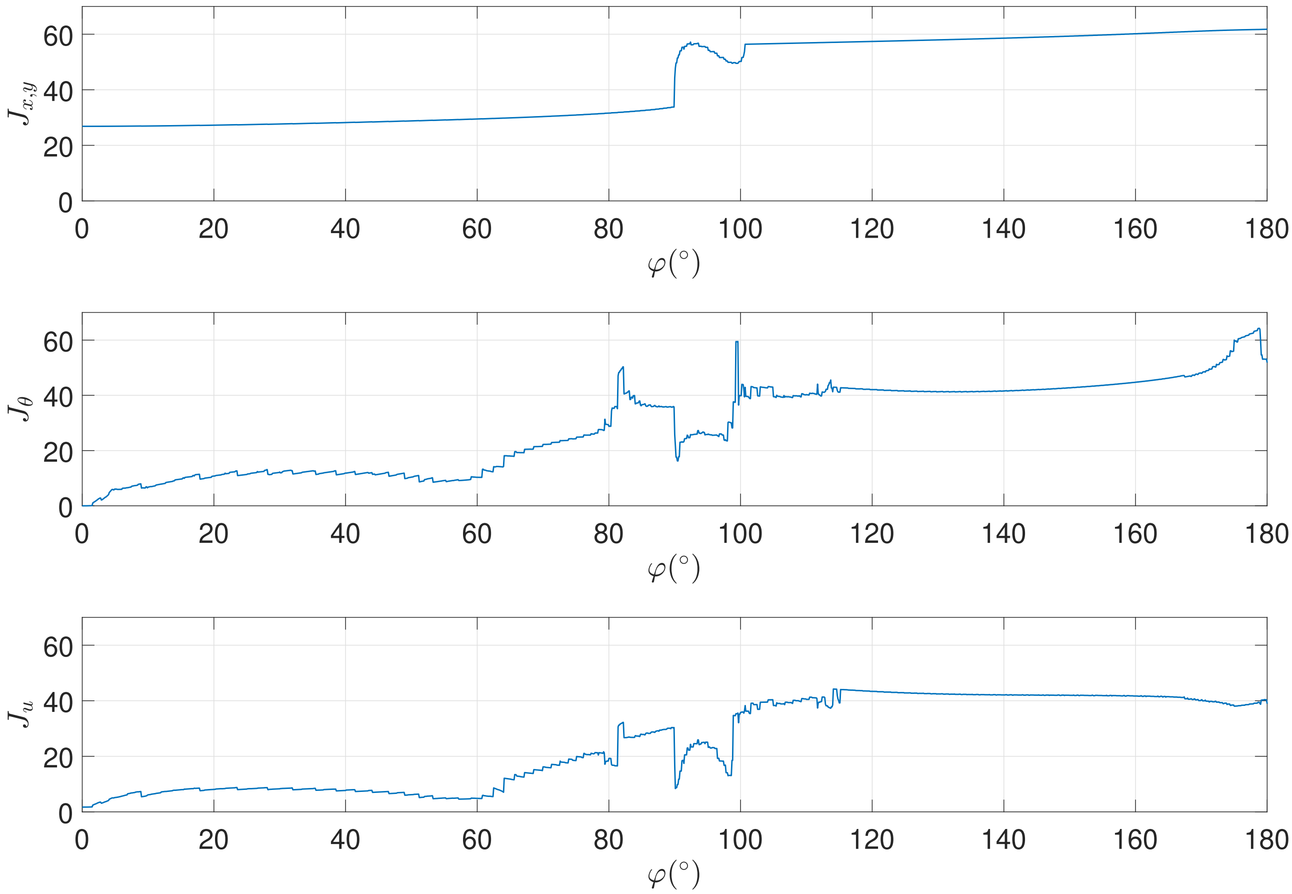


Figure 4-4: A separation of the effect of (top) x and y, (middle) θ and (bottom) u on the cost J.

In Figure 4-4 a constant increase of $J_{x,y}$ can be observed except for the region $90^\circ \leq \varphi \leq 101^\circ$. While for J_θ numerous sudden changes in the value can be observed, thus showing that erratic changes are caused in the orientation when using a lower value of ϵ. To better examine what is causing these sudden changes, in the two following subsections the response of $J_{x,y}$ and J_θ are reviewed respectively. J_u is omitted for review because the graph can be directly related to $J_{x,y}$ and J_θ.

Analysis of HAIC in the x, y dimension.

In the top graph of Figure 4-4, a strange effect of $J_{x,y}$ can be seen starting at $\varphi \geq 90^\circ$. To better analyse this, the trajectory generated for the three values of $\varphi = 89.9$, $\varphi = 90.0$, and $\varphi = 92.5$ are shown in Figure 4-5. These three values are chosen because $\varphi = 89.9$ has a value of $J_{x,y} = 33.8$, while $\varphi = 90.0$ has an increase of 10.4 with a value of $J_{x,y} = 44.2$. The increase of $J_{x,y}$ continuous until $\varphi = 92.5$ with a peak of $J_{x,y} = 56.2$.

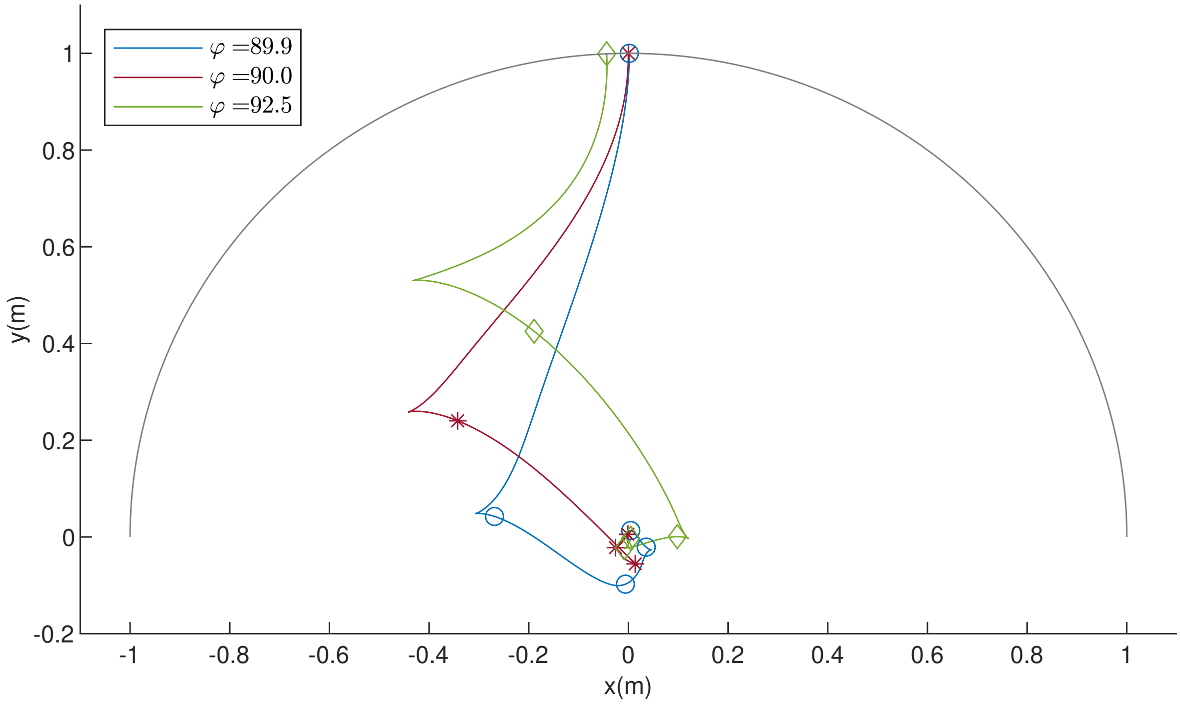


Figure 4-5: Three trajectories generated for starting position corresponding to $\varphi = 89.9$, $\varphi = 90.0$, and $\varphi = 92.5$. The marks correspond to a timestamp at a 10s interval. The first mark corresponds to $t = 0$, the second mark along the trajectory to $t = 10$, the second mark along the trajectory to $t = 20$, etc. The first mark is the one on the grey half-circle.

From Figure 4-5 a difference in trajectory is observable. The moment φ reaches 90°, a two-point reverse turn is initiated earlier in trajectory by the HAIC controller. By creating this turn earlier in the trajectory, the cost function $J_{x,y}$ increases since the turn causes the robot to spend a longer amount of time at a greater distance to the origin. This can be seen when checking the 10-second marks in Figure 4-5 and relating the distance in the $x - y$ plane to each trajectory.

It is important to note that the cost relating to the $x - y$ plane ($J_{x,y}$) and the convergence time are not correlated. The convergence time is highlighted for the three values of $\varphi = 89.9$, $\varphi = 90.0$, and $\varphi = 92.5$ in Figure 4-6. Here it can be seen that the convergence time varies and is more correlated to the overshoot in the $x - y$ plane when moving in on the origin. The convergence time is lowest for $\varphi = 90.0$ which has a minimal overshoot at $t = 20$ as can be seen in Figure 4-5.

A more thorough investigation of what is causing the shift in the two-point reverse turn can be done by looking at the values of $\mu_{r,\theta}$, y and θ. The values of $\mu_{r,\theta}$ and the two states are shown in Figure 4-7. The graphs only contain the first 20 seconds, representing the trajectory from the first until the third marker in Figure 4-5 for the three values of φ.

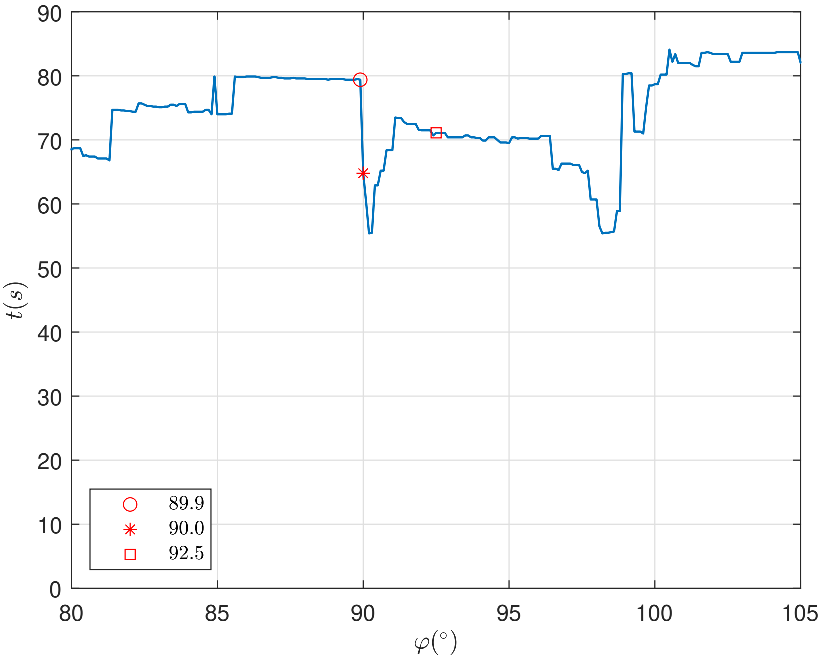


Figure 4-6: Convergence time for values of φ $80°$ until $105°$. The values for φ corresponding to Figure 4-5 are highlighted.

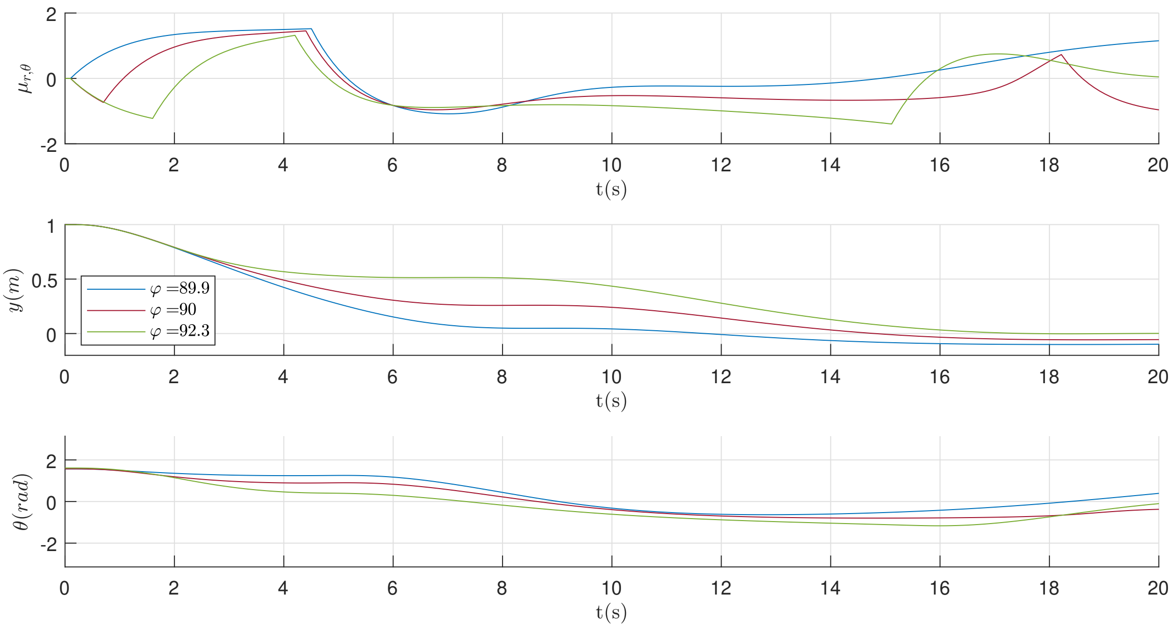


Figure 4-7: The three graphs only contain the first 20 seconds, representing the trajectory from the first until the third mark for the three values of φ. The legend present in the middle graph represents the colour coding for all the graphs.

In Figure 4-7 it is important to note that on the border of $\varphi = 90.0°$ there is an immediate switch for $\mu_{r,\theta}$. The value goes to a negative value for $\varphi \geq 90.0°$ and to a positive for $\varphi < 90.0°$ as can be seen in the top graph of Figure 4-7. This is because $\mu_{r,\theta}$ is updated to have the same value as $arctan(\mu'_y/\mu'_x)$, which the moment $\mu'_x \geq 0$, μ'_y/μ'_x becomes $-\infty$. As stated earlier, μ'_x updates to a value in between $-x$ and $\dot{x}$, similar to μ'_y in between $-y$ and $\dot{y}$.

The tracking of the value $\mu_{r,\theta}$ by θ contains a large error. Only a small adjustment of θ can be seen compared to the dynamics of $\mu_{r,\theta}$. Of course, it is good to note that there was not tuning done to minimize the tracking error. A faster update to the desired value could be obtained by adjusting the hyperparameters regarding the update of μ and/or u.

Referring back to Figure 4-4, a change in value for $J_{x,y}$ can be seen between the values of $\varphi = 92.5$, $\varphi = 99.3$, and $\varphi = 100.7$. Where $\varphi = 92.5$ has a peak value of $J_{x,y} = 56.2$, $\varphi = 99.3$ has the lowest value of $J_{x,y} = 49.8$, and $\varphi = 100.7$ has a value of $J_{x,y} = 56.4$ after which the value for $J_{x,y}$ stays relatively constant. To see what is causing this change, the trajectories for the three values of φ are shown in Figure 4-8.

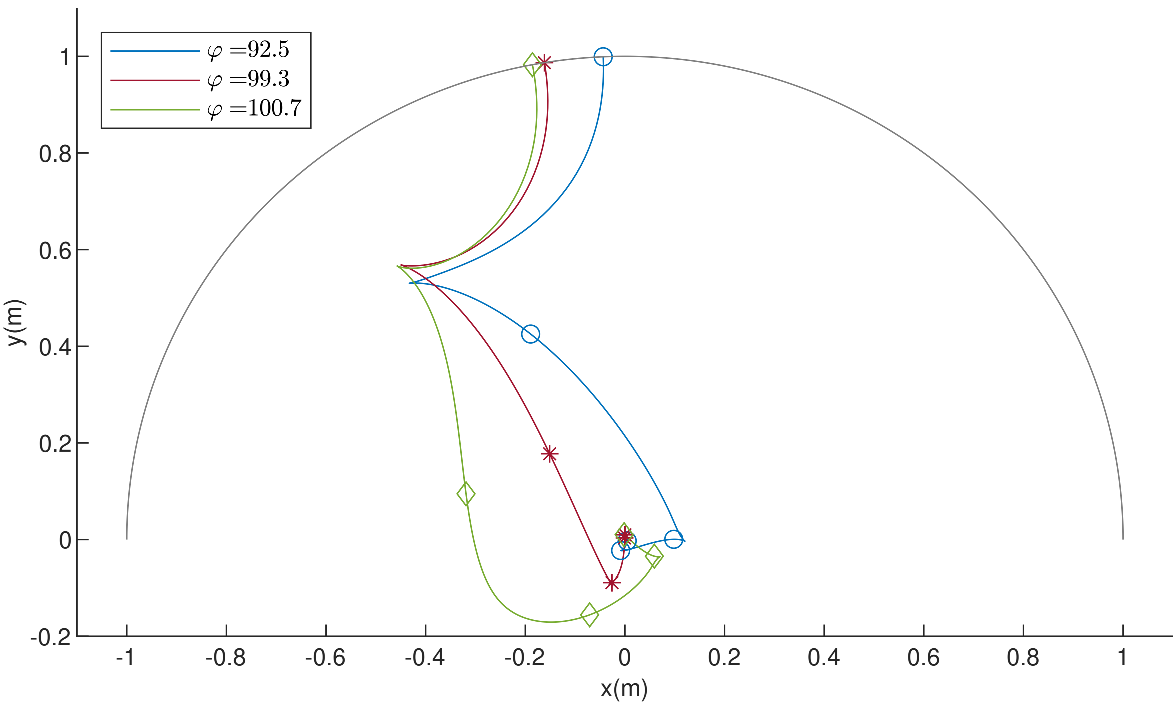


Figure 4-8: Three trajectories generated for starting position corresponding to $\varphi = 92.5$, $\varphi = 99.3$, and $\varphi = 100.7$. The marks correspond to a timestamp at a 10s interval. The first mark corresponds to $t = 0$, the second mark along the trajectory to $t = 10$, the second mark along the trajectory to $t = 20$, etc. The first mark is the one on the grey half-circle.

It is clear that the larger values of $J_{x,y}$ for $\varphi = 92.5$ and $\varphi = 100.7$ can be related to a larger overshoot in the $x - y$ plane compared to the trajectory generated for $\varphi = 99.3$. This shows again that HAIC does not generate an optimal trajectory. However, it does show that HAIC can compensate for any overshoot happening to the system. This also gives an indication that HAIC can easily compensate for disturbances happening to the system.

A question for future research might be how hyperparameters can be obtained such that for different starting states a more constant value of $J_{x,y}$ can be obtained.

Analysis of HAIC in the θ dimension

In Figure 4-4 large variations of J_θ can be seen for small variations of φ. First, this proves that the control of θ is prone to small variations. Whether this can be solved by adjusting the HAIC algorithm is a discussion for future work. In this thesis, a focus is on the origin of the larger variations. To better analyse this, a handful of trajectories in the range of $80° \leq \varphi \leq 105°$ are highlighted.

It was shown in Figure 4-2 that even for a higher value of ϵ these variations were still applicable. In Figure 4-9 this region and the response J_θ is shown. Three groups (labelled by colour) which each contain three values of φ (labelled by varying marks) are shown. These three groups are chosen because of the large change of J_θ and will be discussed in this subsection.

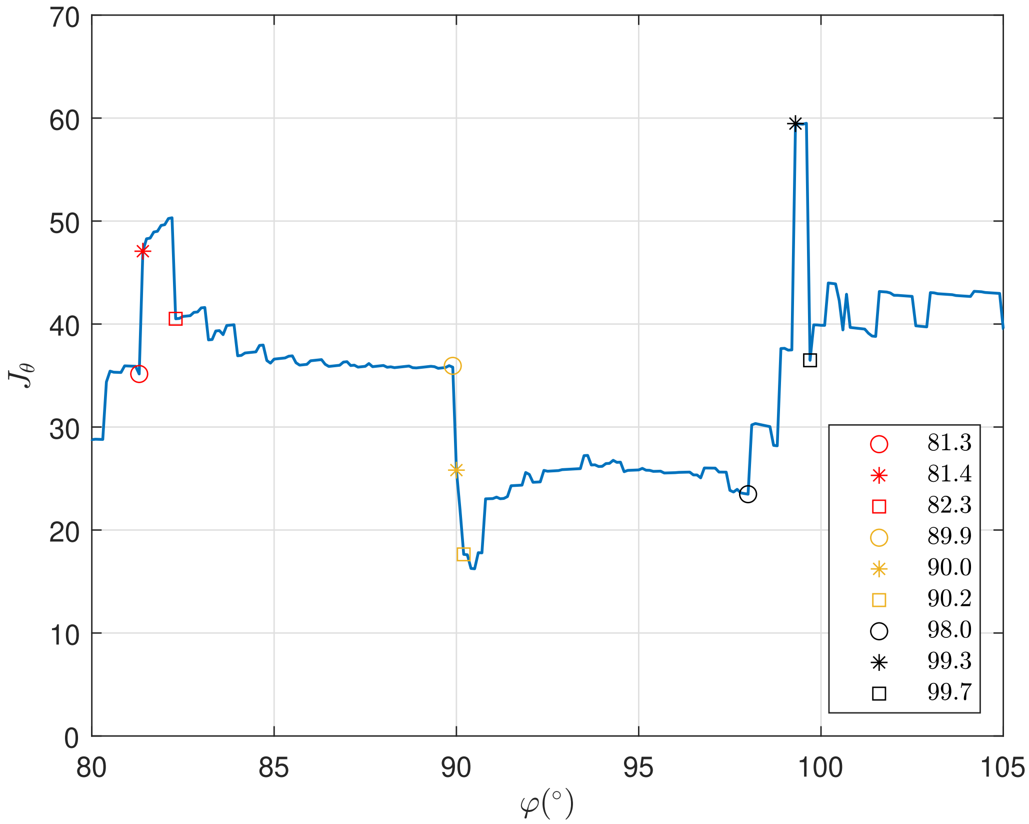


Figure 4-9: This figure is a zoomed version of the middle graph in Figure 4-4. Trajectories for the three groups, Red, Yellow, and Black, can be seen in Figure 4-10, Figure 4-12, and Figure 4-13 respectively. The values in the legend represent the values for φ in degerees.

The first group discussed is that of $\varphi = 81.3$, $\varphi = 81.4$, and $\varphi = 82.3$. Within one iteration the value for J_θ goes from 35.2 to 47.1 for respectively $\varphi = 81.3$ and $\varphi = 81.4$. After which J_θ goes from 50.3 to 40.5 for $\varphi = 82.2$ and $\varphi = 82.3$.

Trajectories corresponding to $\varphi = 81.3$, $\varphi = 81.4$, and $\varphi = 82.3$ are shown in Figure 4-10. It can be seen that on a large scale almost no difference is noticeable. But when zoomed in around the origin, a small three-point turn is made at the first mark (fourth mark in the full-size graph) is made for the trajectory of $\varphi = 81.4$. This is caused by the overshoot in the $x-y$ plane. $\varphi = 81.3$ does not contain the additional turn, and $\varphi = 82.3$ does not contain a three-point turn.

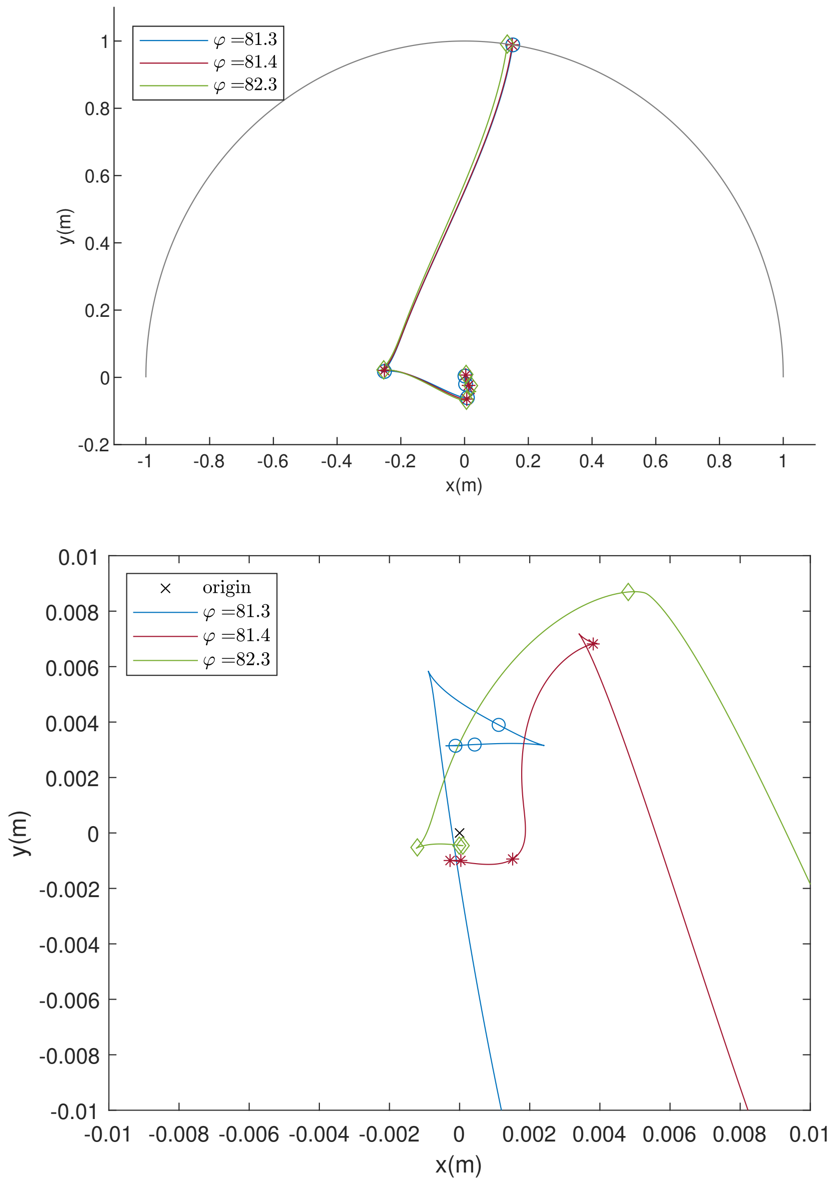


Figure 4-10: Three trajectories generated for starting position corresponding to $\varphi = 81.3$, $\varphi = 81.4$, and $\varphi = 82.3$. The marks correspond to a timestamp at a 10s interval. In the top figure, the first mark corresponds to $t = 0$, the second mark along the trajectory to $t = 10$, the third mark along the trajectory to $t = 20$, etc. The first mark is the one on the grey half-circle. In the bottom figure, the line coming from outside the graph is correlated to the lowest value of t. The first mark corresponds to $t = 30s$, the second to $t = 40s$.

The question is whether the overshoot is caused by HAIC updating its internal beliefs and control input not fast enough, or by HAIC giving a non-optimal control input. By checking the evolution of the internal beliefs of the controller for $\varphi = 82.4$ an insight can be given on what the issue might be. The controller works with 11 internal beliefs of which 7 are updated to observations. These 7 beliefs are plotted for $\varphi = 81.4$ in Figure 4-11 together with their corresponding observations. The belief for θ also has the belief of the $\mu_{r,\theta}$ plotted since this is the reference for this belief.

From Figure 4-11 it can be seen that the beliefs track the observations with an error. This is a common issue with AIC in general and referred to as biased state estimation [29]. The belief is updated towards both the goal and the observation. Although this bias might not be large for a single belief, what happens is that the bias continues to give misinformation to other beliefs. A bias for μ_x causes a biased update for μ'_x since it makes use of μ_x for the update. With the addition of HAIC and thus an extra layer of beliefs, which in turn are partly dependent on μ'_x, the bias will continue to grow. Thus, likely causing an overshoot. The goal of this thesis is not to make HAIC the state-of-the-art controller, therefore no further investigation for a solution to this problem will be discussed. To make HAIC a competitive controller future work can be done to prevent biased state estimation and the continuous growth of bias for every belief.

When the other two groups in Figure 4-9 are plotted in Figure 4-12 and Figure 4-13 a similar overshoot can be seen. In Figure 4-12 an overshoot in $x - y$ plane at 20 seconds can be observed for the trajectory corresponding to $\varphi = 89.9$. After which the robot needs to correct with an *angle* of 90°. This overshoot is not present when a two-point turn is initiated earlier in the trajectory. In Figure 4-13 two overshoots in the $x - y$ plane are present for $\varphi = 99.3$ which had the highest value of J_θ for any φ. In the top graph, there is an overshoot at 20 seconds, and in the bottom graph at 30 seconds another overshoot which causes the controller to set θ to a $-90°$ and $90°$ angle respectively.

Giving two additional examples that the overshoot caused by the controller is giving the large values of J. The fact that these high values of J and the corresponding overshoots are happening on rare occasions show the robustness of an untuned HAIC controller.

The examples discussed both for $J_{x,y}$ and J_θ show that inconsistency of HAIC is mainly caused due to overshoot caused by the control. Still HAIC can always converge a nonholonomic system from any starting point. It seems that when HAIC is untuned it is prone to overshoot and not obtain smooth control trajectories. Future work could therefore focus on improving the control of HAIC such that smooth trajectories are obtained.

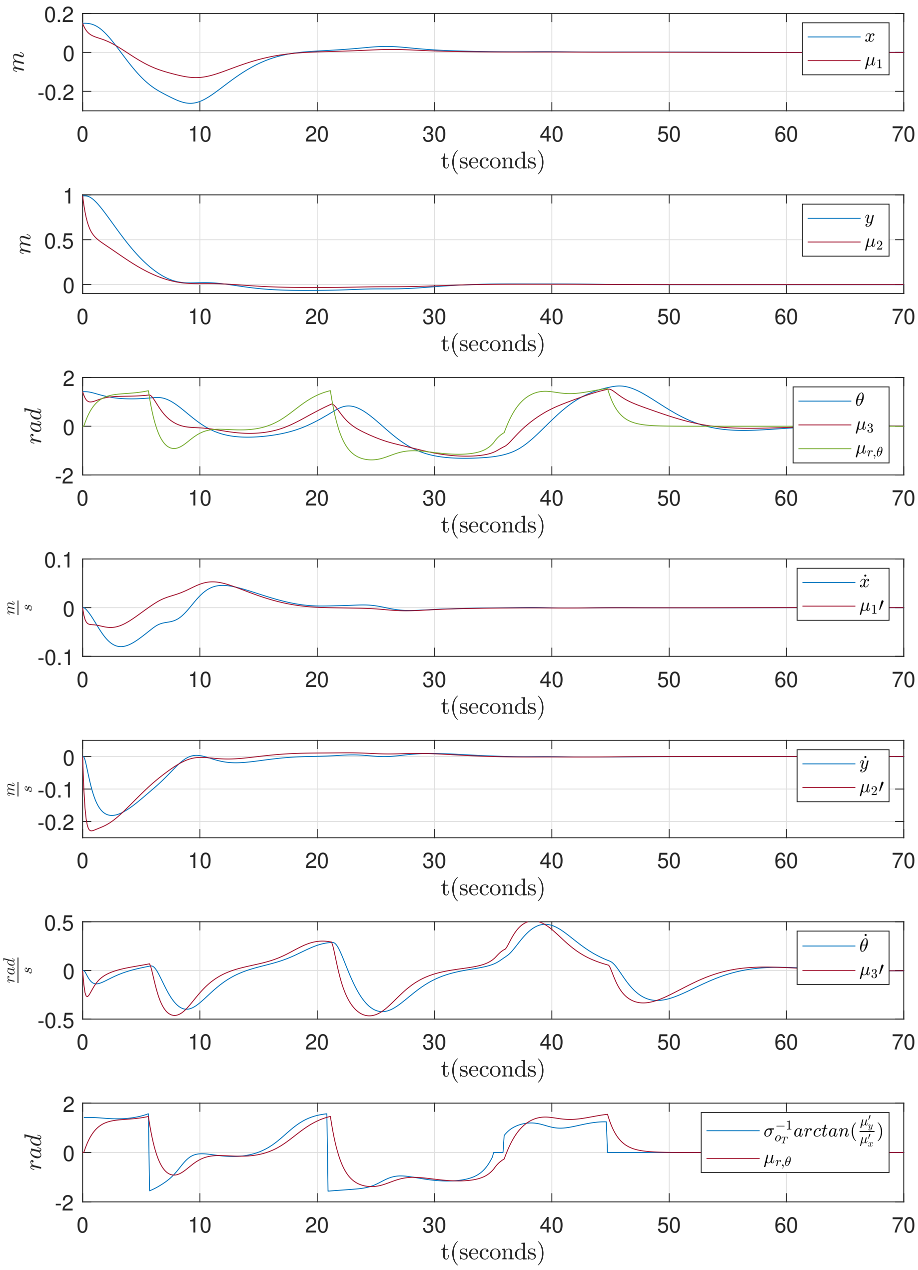


Figure 4-11: The 7 beliefs for the trajectory of $\varphi = 82.4$ are plotted. All the blue lines are the observations for the beliefs, when there is a green line it is the reference for the belief. If there is no green line, the reference is zero. (For grey scale viewers, blue lines are the first legend entry, and the optional green lines are the third legend entry)

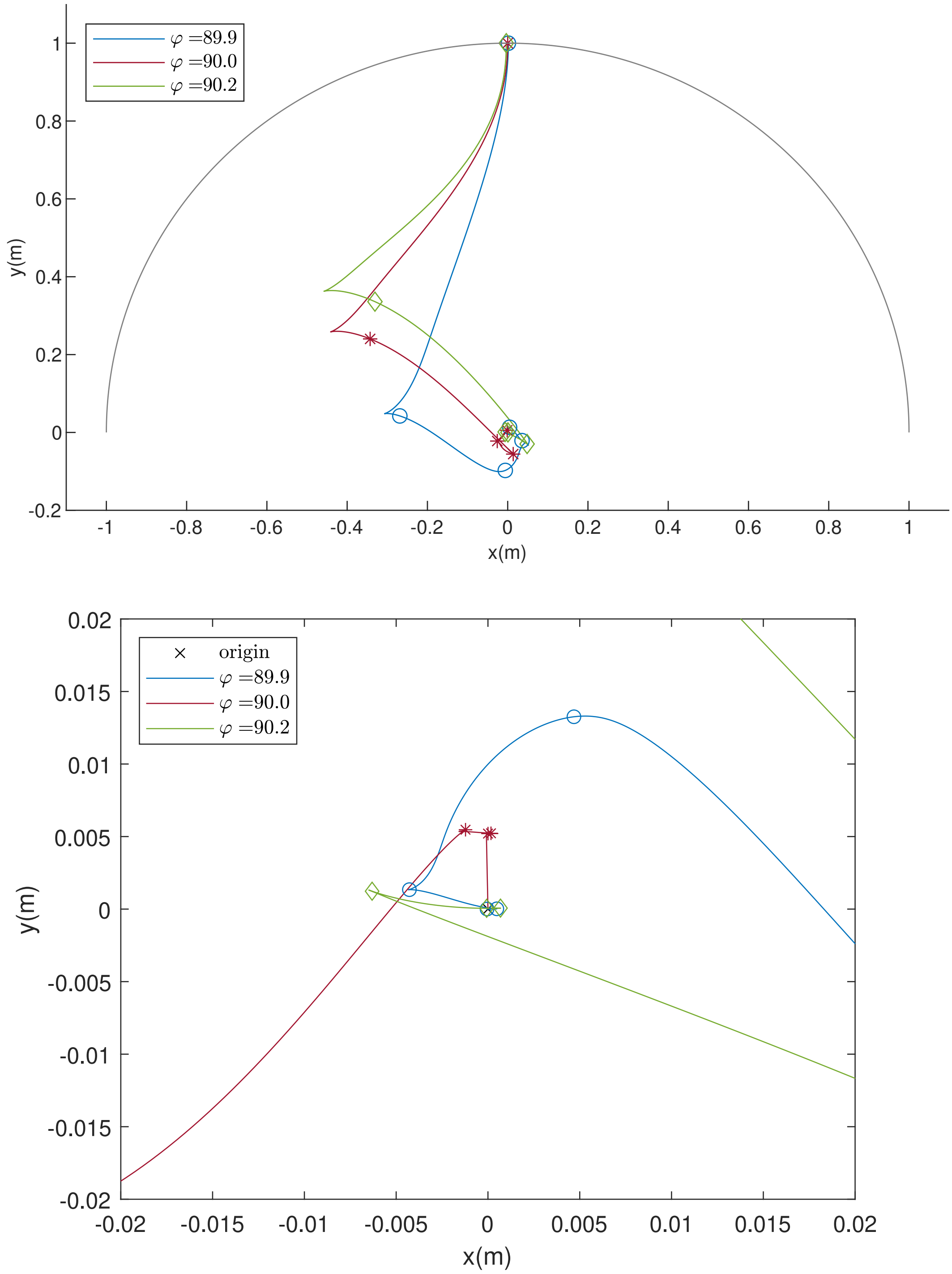


Figure 4-12: Three trajectories generated for starting position corresponding to $\varphi = 89.9$, $\varphi = 90.0$, and $\varphi = 90.2$. The marks correspond to a timestamp at a 10s interval. In the top figure, the first mark corresponds to $t = 0$, the second mark along the trajectory to $t = 10$, the third mark along the trajectory to $t = 20$, etc. The first mark is the one on the grey half-circle. In the bottom figure, the line coming from outside the graph is correlated to the lowest value of t. The first mark corresponds to $t = 30s$, the second to $t = 30s$.

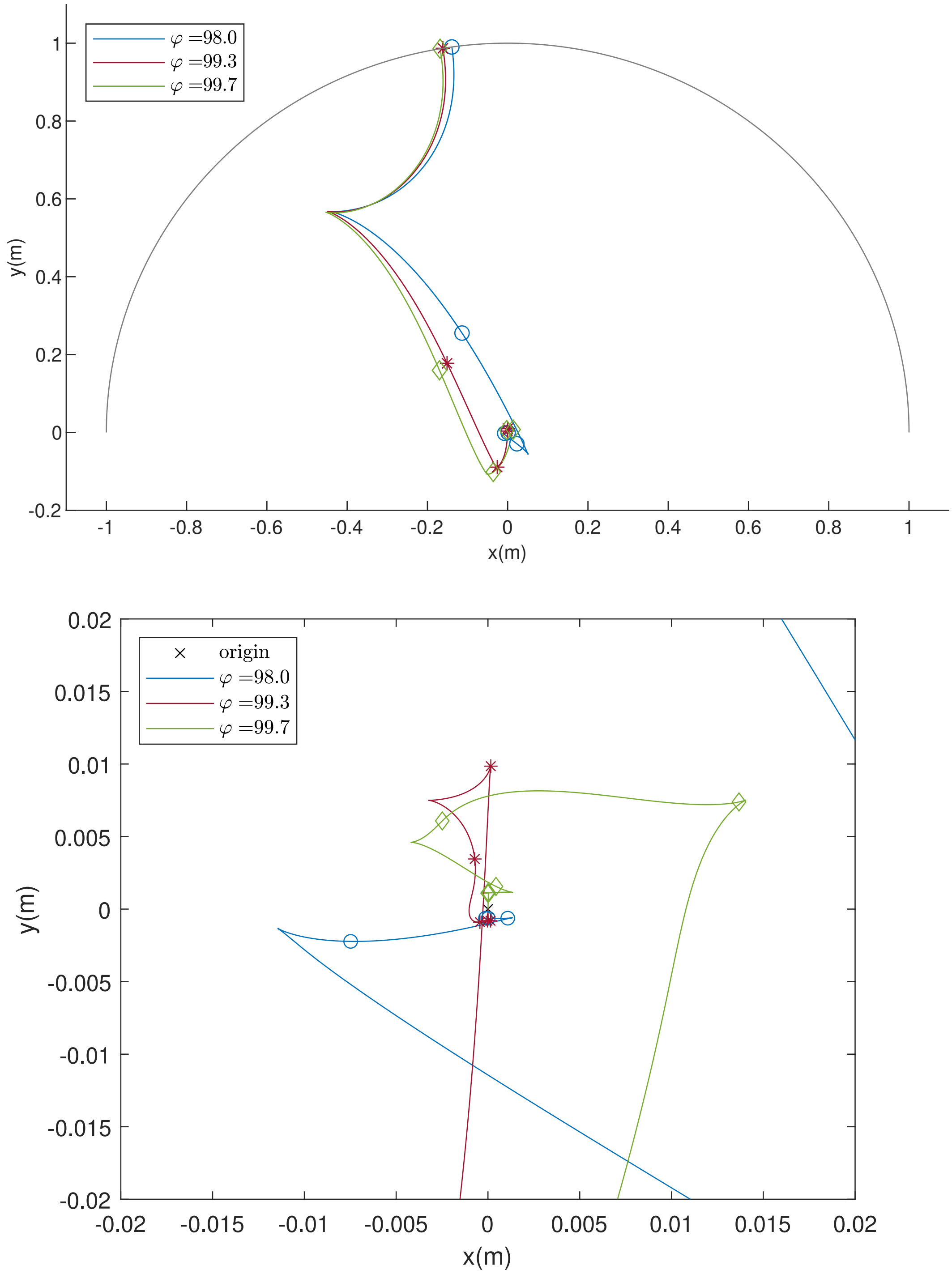


Figure 4-13: Three trajectories generated for starting position corresponding to $\varphi = 98.0$, $\varphi = 99.3$, and $\varphi = 99.7$. The marks correspond to a timestamp at a 10s interval. In the top figure, the first mark corresponds to $t = 0$, the second mark along the trajectory to $t = 10$, the third mark along the trajectory to $t = 20$, etc. The first mark is the one on the grey half-circle. In the bottom figure, the line coming from outside the graph is correlated to the lowest value of t. The first mark corresponds to $t = 30s$, the second to $t = 30s$.

4-2 Noise

In section 3-4, the hyperparameter ϵ is introduced. It is a newly introduced hyperparameter that is required to be tuned accordingly such that the system can converge to a reference point. In this section, an example is shown of why it is required to tune accordingly and what happens when it is not done.

The beliefs μ are updated according to observations. These observations can be direct measurements that can contain different forms of noise. Because the dynamics of μ are directly related to the observations, μ can obtain a similar noisy behaviour as the observations. To show this, a trivial control problem of a two-wheeled mobile robot in simulation is shown. The robot is controlled to the origin while starting at the set point $z = [-\frac{1}{2}\sqrt{3} \;\; -\frac{1}{2}\sqrt{3} \;\; 0]^T$. Additive white noise is added to the observations in the form of:

$$\begin{aligned} o &= \begin{bmatrix} x \\ y \\ \theta \end{bmatrix} + w, \quad w \sim \mathcal{N}(0, \Sigma_n) \\ \Sigma_n &= \begin{bmatrix} \sigma_n & 0 & 0 \\ 0 & \sigma_n & 0 \\ 0 & 0 & \sigma_n \end{bmatrix} \end{aligned} \tag{4-5}$$

For this example, the three values of 10^{-5}, 10^{-4}, and 10^{-3} are used for the variance σ_n. Note that Σ_o will not be adjusted accordingly and will still be kept as the identity matrix. The reason for this is to show that ϵ can tuned without the need of knowing the actual noise levels.

In the coming four figures the response of a selected number of states and beliefs is shown for the control of the robot. The reason for these four graphs is to show the convergence of θ which will suffer when ϵ is not correctly chosen. The values for $\sqrt{(\mu'_x)^2 + (\mu'_y)^2}$ and ϵ are visible such that the margins for the binary switch are visible. $\sigma^{-1}_{\mu_{r,\theta}}$ and $\sigma^{-1}_{o_T}$ are displayed to show the switching of the values. And last, the value of $c(t)$ on a logarithmic scale is shown to represent the convergence of the system.

In Figure 4-14 the graphs are plotted for $\sigma_n = 10^{-5}$ and $\epsilon = 10^{-3}$. What can be seen is θ converging, as well as the value of $c(t)$. $\sqrt{(\mu_x')^2 + (\mu_y')^2}$ is diving underneath the value of ϵ and stays there. Because of this, we know the value for ϵ is set correctly. There are only a handful of switches happening for the values of $\sigma_{\mu_{r,\theta}}$ and σ_{o_T} at around $t = 30s$. It is also important to note that even though the switches are happening, both $\sqrt{(\mu_x')^2 + (\mu_y')^2}$ and $c(t)$ are still reducing in value. The graph of $c(t)$ jumps further underneath a value of 10^{-5} but for easy comparison the axis is kept at these values.

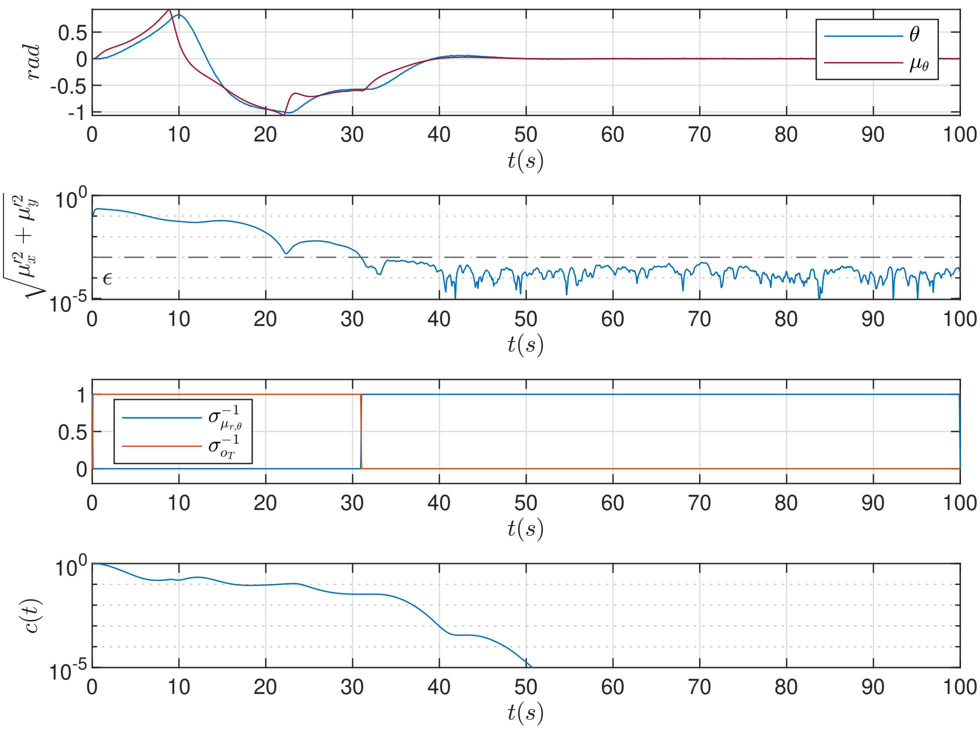


Figure 4-14: Various states and beliefs are shown when observations contain white noise with variance $\sigma_n = 10^{-5}$. The first graph contains the state θ and the corresponding belief. The second graph contains the value which determines the value for $\sigma_{\mu_{r,\theta}}$ and σ_{o_T} and hyperparameter value $\epsilon = 10^{-3}$ represented by the dotted line. The third graph represents the values for $\sigma^{-1}_{\mu_{r,\theta}}$ and $\sigma^{-1}_{o_T}$ which switch based on the second graph. The fourth graph shows the cost at each time step.

When increasing the value of σ_n to 10^{-4}, a response of the controlled system is obtained as in Figure 4-15. For the value of θ some small disturbances can be seen when converged to zero at around $t = 50s$. When looking at the graph showing the value for $\sqrt{(\mu_x')^2 + (\mu_y')^2}$, it can be seen that the value is kept just under the value of ϵ. Though occasionally, $\sqrt{(\mu_x')^2 + (\mu_y')^2}$ reaches a value higher than ϵ which is related to the switching of $\sigma_{\mu_{r,\theta}}^{-1}$ and $\sigma_{o_T}^{-1}$. The non-convergence of the system is also represented in the non-converged value of $c(t)$. Overall, it seems like θ is unable to properly converge and contains some small disturbances.

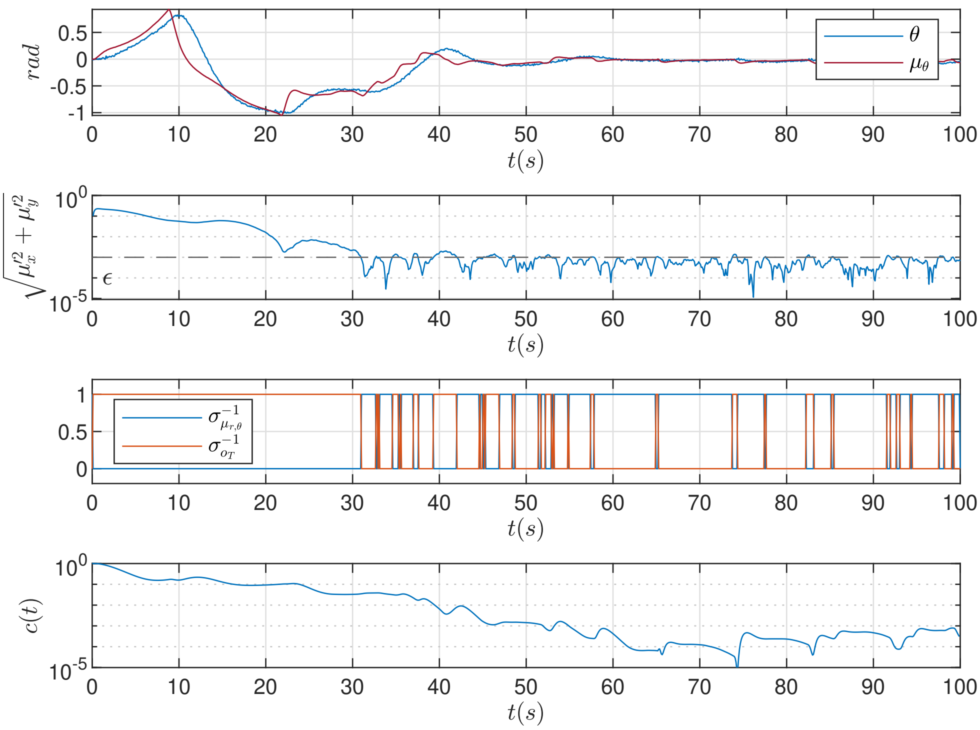


Figure 4-15: Various states and beliefs are shown when observations contain white noise with variance $\sigma_n = 10^{-4}$. The first graph contains the state θ and the corresponding belief. The second graph contains the value which determines the value for $\sigma_{\mu_{r,\theta}}$ and σ_{o_T} and hyperparameter value $\epsilon = 10^{-3}$ represented by the dotted line. The third graph represents the values for $\sigma_{\mu_{r,\theta}}^{-1}$ and $\sigma_{o_T}^{-1}$ which switch based on the second graph. The fourth graph shows the cost at each time step.

In Figure 4-15 the response of the system is shown when the value of $\sqrt{(\mu_x')^2 + (\mu_y')^2}$ can reach lower than the value of ϵ. In Figure 4-16 the value for σ_n is increased to 10^{-3} and the system operates with a value for $\sqrt{(\mu_x')^2 + (\mu_y')^2}$ just above the value of ϵ.

It is clear that ϵ is set incorrectly because of the inability to converge θ to a constant value as visible in the top graph of Figure 4-16. The value for ϵ is too low, which causes $\sigma_{\mu_{r,\theta}}^{-1}$ to never obtain a value of 1 for a long duration. This causes HAIC to control θ to a value of $arctan(\mu_y'/\mu_x')$ instead of zero when x and y value are close to zero.

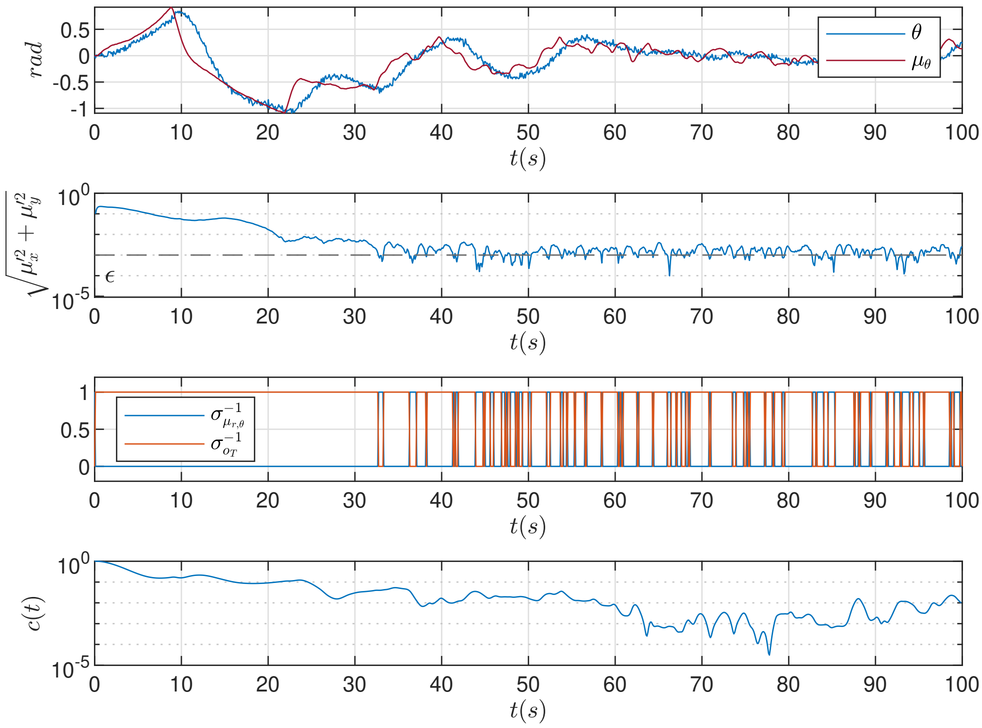


Figure 4-16: Various states and beliefs are shown when observations contain white noise with variance $\sigma_n = 10^{-3}$. The first graph contains the state θ and the corresponding belief. The second graph the contains the value which determines the value for $\sigma_{\mu_{r,\theta}}$ and σ_{o_T} and hyperparameter value $\epsilon = 10^{-3}$ represented by the dotted line. The third graph represents the values for $\sigma_{\mu_{r,\theta}}^{-1}$ and $\sigma_{o_T}^{-1}$ which switch based on the second graph. The fourth graph shows the cost at each time step.

A simple adjustment of ϵ to deal with noisy observations with a variance of 10^{-3} is the only thing necessary such that HAIC can converge the system again. In Figure 4-17, ϵ is increased to a value of 10^{-2} as can be seen in the second graph. The value of $\sqrt{(\mu_x')^2 + (\mu_y')^2}$ stays underneath this value. A single switch of the values of $\sigma_{\mu_{r,\theta}}^{-1}$ and $\sigma_{o_T}^{-1}$ can be observed. Both θ and $c(t)$ now converge to a constant value.

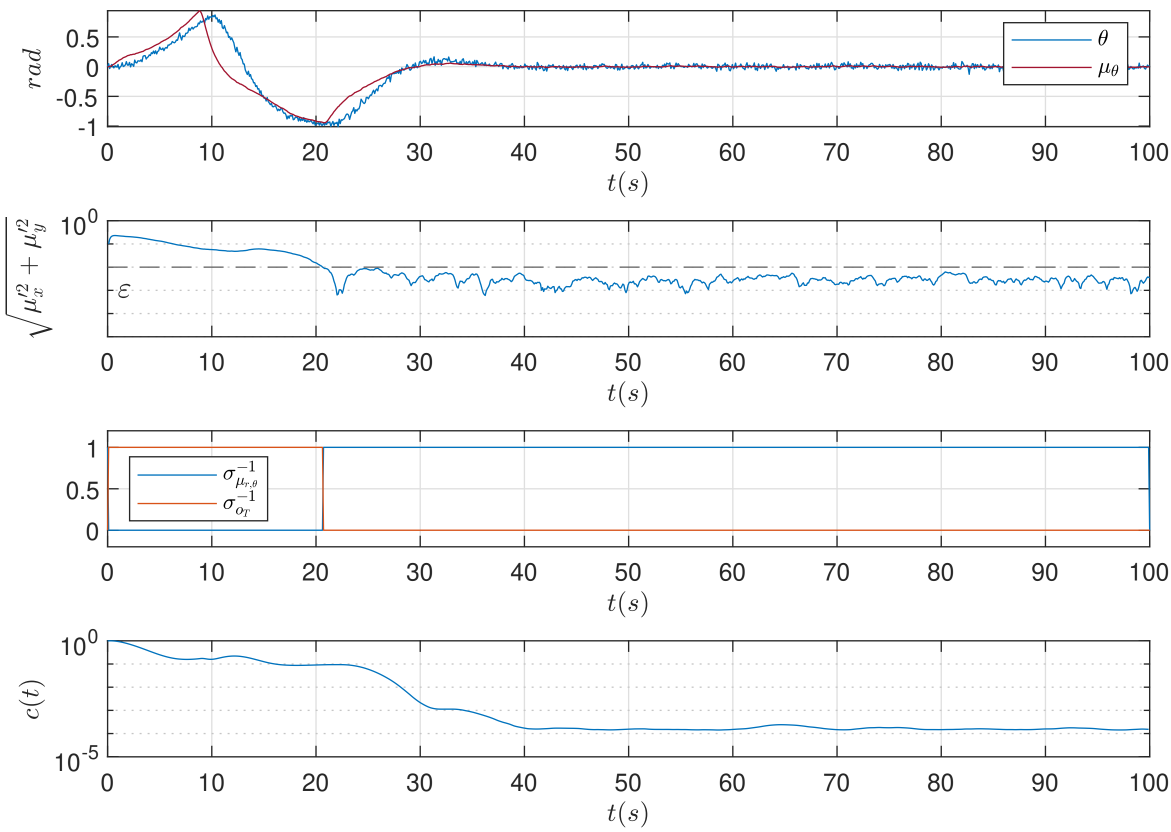


Figure 4-17: Various states and beliefs are shown when observations contain white noise with variance $\sigma_n = 10^{-3}$. The first graph contains the state θ and the corresponding belief. The second graph the contains the value which determines the value for $\sigma_{\mu_{r,\theta}}$ and σ_{o_T} and hyperparameter value $\epsilon = 10^{-2}$ represented by the dotted line. The third graph represents the values for $\sigma_{\mu_{r,\theta}}$ and σ_{o_T} which switch based on the second graph. The fourth graph shows the cost for each time step.

These simulations show that ϵ is prone to noise contained in the observations used for HAIC. It is highly likely that other disturbances, such as incorrect dynamics or external disturbances are also of influence and require the adapting of the value ϵ. The simulations also show that when the noise of the observations is not known, the response of the system can show that the value for ϵ is incorrect. When the system does reach the origin but is constantly rotating at its location, it is likely that ϵ is set to low.

As stated in previous research, tuning of parameters like that of the hyperparameter Σ_o can make sure the beliefs transfer less of the noise contained in the observations [1]. It is highly likely that adaptation of Σ_o can also increase the change of convergence. However, this thesis does not dive into the tuning of the other parameters of AIC and HAIC. A possibility would be to take this into account in future work.

4-3 Conclusion

From a varying range of starting points, HAIC can control a nonholonomic system to a three-dimensional reference state. It is obvious that HAIC does not obtain an optimal control input or trajectory as can be seen from the comparison with MPC and the example trajectories discussed in this chapter. When applying HAIC it is important to note that a larger ϵ allows for faster convergence of the system, but does increase the error to the set-point when converged.

The first simulation indicates that the combination of AIC and Hierarchical Active Inference (HAI) can be combined as described according to section 3-4 to successfully control a nonholonomic system. Thus, having answered the second research question as stated in section 1-2.

Small variations in starting state can cause large variations in trajectories, especially when operating around the border of the $arctan(\cdot)$ function. The reference given to the bottom layer using $\mu_{r,\theta}$ is not tracked strictly, and an error is noticeable when compared to the value of μ_θ. This is likely because of the biased estimation, when solved it could make HAIC a more competitive controller. A step in this direction has already been done for AIC which was referred to as Unbiased AIC[29].

The effect of the hyperparameter ϵ of HAIC was investigated when observations contained additive white noise. As expected, tuning was required to account for the noise such that convergence can be obtained. When noise levels of observations are unknown, the response of the system can give an indication whether ϵ is set correctly. Where the ability to converge the robot in x and y dimension but not in θ dimension indicates that ϵ is set too low. More research could be done regarding different colours of noise and how to deal with the inability to converge online instead of manual tuning.

The simulations prove that ϵ is required to be set to a correct value. Answering the question "Which parameters have to be set in addition to AIC, and what influences their values, such that convergence can be obtained when using HAIC for the control of a two-wheeled mobile robot?". The noise of the observations, transferred to the beliefs $\tilde{\mu}$ give an indication of the required value. Where a higher noise value in the observations requires a higher value for ϵ.

Throughout the simulations it was clear that control of HAIC can be improved. Creating HAIC such that it can control a system smooth and fast is out of the scope of this thesis. Future work could be performed to research how to obtain smooth and fast control with HAIC. Possibilities could also contain aspect of adaptive and self tuning control.

Chapter 5

Physical Experiment

In this chapter, Hierarchical Active Inference Control (HAIC) is applied to a physical two-wheeled differential-drive mobile robot made by Avular, called Cody [35]. The experiment setup is clarified with the tools used for the tests done with the robot. This includes an explanation of the robot and the hardware and software used for the experiment. After which the results are discussed for two different values of ϵ. From the results, a conclusion will be drawn and a comparison to the simulations can be made.

5-1 Experiment Setup

The experiment exists of a single two-wheeled robot in an arena, which is a square of three by three meters. Within this arena, five positions with orientation are given as reference points for the HAIC controller.

In chapter 2, it is stated that the robot is always controlled to the origin. The five positions with orientation are given in the global frame, this is indicated by the addition of the superscript ^{g}. However, the results will show a robot which is controlled to the origin, the control is done in the reference frame (all states without superscript are in the reference frame). The references are given in the global coordinate frame as:

$$z^g_{r,1} = \begin{bmatrix} -0.5 \\ -0.5 \\ \pi \end{bmatrix}, \quad z^g_{r,2} = \begin{bmatrix} 0.5 \\ 0.5 \\ 0 \end{bmatrix}, \quad z^g_{r,3} = \begin{bmatrix} 0.5 \\ -0.5 \\ \frac{\pi}{2} \end{bmatrix}, \quad z^g_{r,4} = \begin{bmatrix} -0.5 \\ 0.5 \\ -\frac{\pi}{2} \end{bmatrix}, \quad z^g_{r,5} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{5-1}$$

Throughout the experiment HAIC will control Cody from $z^g_{r,5}$ to $z^g_{r,1}$, from $z^g_{r,1}$ to $z^g_{r,2}$, from point$z^g_{r,2}$ to point $z^g_{r,3}$, ..., until it reaches $z^g_{r,5}$, and it goes back again to $z^g_{r,1}$, this cycle is repeated to obtain samples. Ten samples are obtained, such that ten similar trajectories are obtained for the five different reference states. The states in their global dimensions are represented in Figure 5-1.

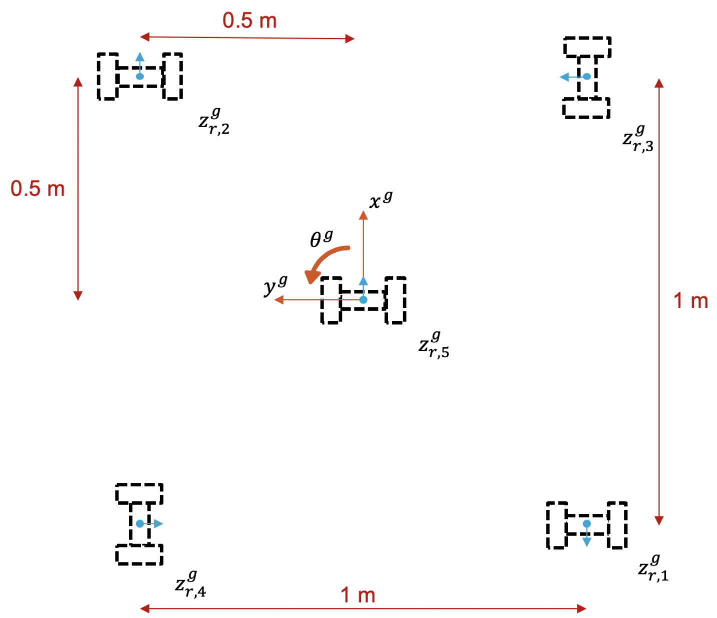


Figure 5-1: In this figure, a representation of the center of the arena is shown which includes all the reference points used for the experiment. The blue dots with arrows inside the dotted robots correspond to the reference points $z^g_{r,i}$. $z^g_{r,5}$ indicates the center of the arena, and the rest of the figure represents a square of a size of 1 by 1 meter surrounding the area.

From these five different reference points, trajectories will be obtained with can be grouped into five similar trajectories. These trajectories will be indicated as: trajectory one (traj-1) going from $z^g_{r,5}$ to $z^g_{r,1}$, (traj-2) going from $z^g_{r,1}$ to $z^g_{r,2}$ ...., until trajectory five (traj-5). In Figure 5-2, the trajectories are shown from the start position to the reference position. The notation of traj-1, traj-2, ..., traj-5 for the trajectories is used such that in the section 5-2 a differentiation can be made between trajectories, even though all trajectories go to the origin according to the reference frame used by HAIC as explained in chapter 2.

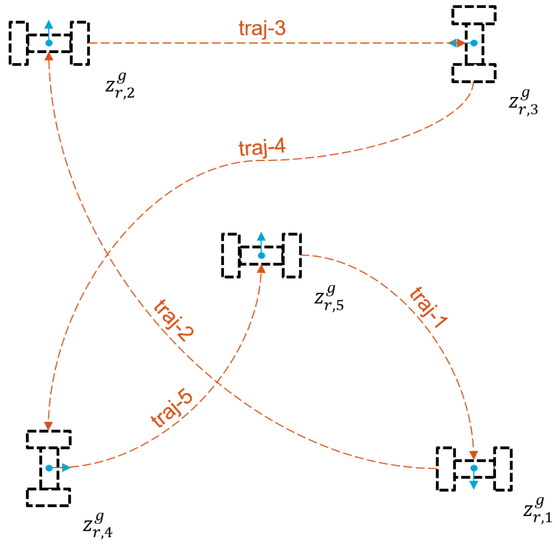


Figure 5-2: In this figure, the trajectories from traj-1 to traj-5 are shown. Note that these trajectories do not correspond to any obtained trajectory from the experiment but are just an indication of from where to where the robot will travel.

The Mobile Robot

Cody is a two-wheeled differential-drive mobile robot from the robotics company Avular [35]. Cody consists of four contact points to the ground. Two large actuated wheels on the side of the robot, and two omnidirectional free roller joints. The omnidirectional rollers are not actuated and are only there to prevent any tipping-over of the robot, one is placed in front of the robot and the other behind the robot. In Figure 5-3 Cody can be seen from two different perspectives. The robot makes use of a Raspberry Pi for actuator control. On this Raspberry Pi, the control is implemented using Ubuntu 22.04 and ROS 2.

Figure 5-3: Two pictures showing the mobile robot Cody. The big wheel visible in both pictures is the actuated differential drive wheel. The free roller wheels are the synapse like

The design of Cody is mainly 3D printed where wheel and joint resistance were not a main factor for the design decisions. The robot therefore has possible wheel slip, the wheels are 3D printed where each wheel has three small rubber bands. The two large actuated wheels can therefore lose grip on the ground and turn without moving the robot. The two small omni-directional wheels may get blocked because of dirt building up in the joint holds. Resistance present in the large wheels also causes the wheels to not move when small amounts of control input are applied. Overall, these can cause disturbances to the system during operation, but they are not considered in the model. In this thesis, these disturbances are not identified, but they are present throughout the experiment.

Accurate Observation

To test the functionality of HAIC, a motion track system is used to obtain a precise position. For this setup, an OptiTrack Flex camera system is used with 6 evenly spread cameras. The trackable nodes defining the rigid body of the robot can be seen on top of the robot in Figure 5-3 (the light grey balls). A laptop is used running Motive OptiTrack and feeding the data over Wi-Fi to the Raspberry Pi. The complete setup can be seen in Figure 5-4.

Figure 5-4: Setup used to do experiments. In the top region of the picture, red cameras can be seen. These are used for obtaining an accurate location of the robot. Two laptops are used. The left running Motive OptiTrack to send data to the Raspberry Pi. The right uses an SSH connection to access the terminal on the Raspberry Pi. The robot is in the arena and currently positioned on $z^g_{r,3}$.

5-2 Experiment Results

To test capabilities HAIC for a physical mobile robot, tests described in section 5-1 have been done for two values of ϵ. All other hyperparameters have been set to 1 or the identity matrix as stated in section 3-2. The trajectories (traj-1, ..., traj-5) have been done 10 times for each value of ϵ totalling 100 collected trajectory samples. The values for ϵ have been set at 0.05 and 0.005 to confirm the requirement of setting the value correctly as concluded in chapter 4.

The experiments are shown in box plots. The results for each ϵ are grouped in trajectory, and grouped for absolute converged error in dimensions x, y, and θ. The reason for grouping in trajectory is showcasing that for different starting points, a different error is on average obtained. The reason for grouping in absolute error per dimension, is to show a higher error value in certain dimensions. Where a higher error in the y dimension is obtained when ϵ is set correctly and a higher error in the θ dimension when ϵ is set incorrectly.

The box plots can be read as follows: the box indicates the 25th percentile (bottom) and 75th percentile (top) of the data, the bar inside the box indicates the median of the data, the bars outside the box indicate the minimum and maximum of the data excluding outliers. A data point is considered an outlier when it is outside the range of $1.5 \times IQR$ from the bottom or the top of the box, where the Interquartile Range (IQR) is the distance between the bottom and top of the box [36].

In Figure 5-5, the box plots can be seen for the experiments done with a value of $\epsilon = 0.05$. It can be seen that for all 50 samples the robot converges to the desired set point due to the errors of x and θ being very low. A difference in final cost can be observed for different trajectories in the left figures. Confirming that just like the simulations, different starting points cause a variation in final error. The box plots showing the absolute error of the states of the converged confirm the larger error in the y dimension. This stated in chapter 3 and shown in chapter 4 as a characteristic of HAIC and based on the value of ϵ.

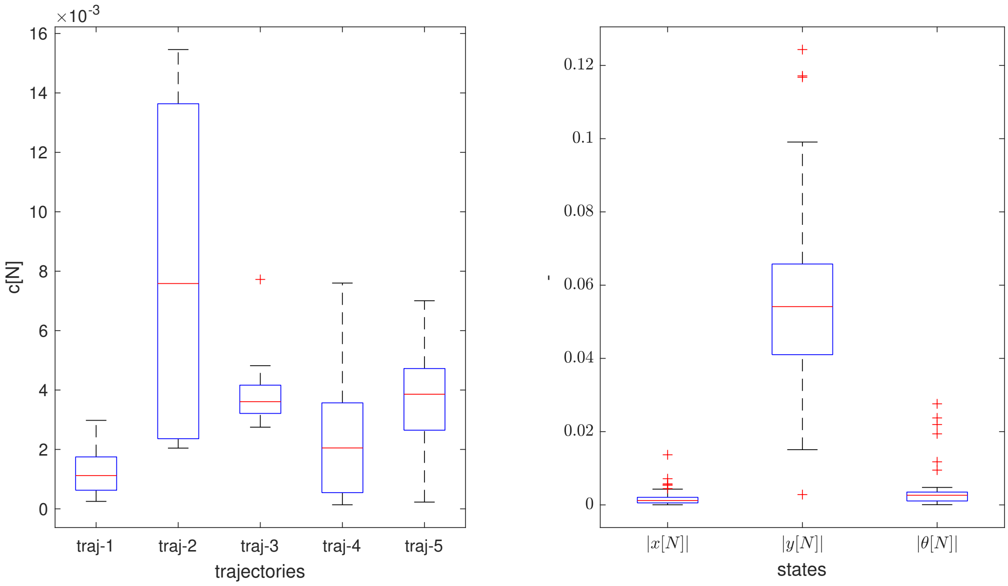


Figure 5-5: Data samples of the converged mobile robot for the value of $\epsilon = 0.05$. (Left) Five box plots of data are shown corresponding to the five trajectories elaborated in section 5-1, each box plot contains 10 data samples. (Right) Three box plots containing data corresponding to the three position states of the converged robot, each box plot contains 50 data samples.

In Figure 5-6, one sample for each trajectory is shown including a top-down photo of a time stamp of the robot. As concluded from the previous box plot, it can be seen that indeed an error occurs perpendicular to the blue arrows. This is the y dimension of the reference frame belonging to the blue arrow, and the frame the robot is being controlled in at that moment.

Just like in the simulations, in the sample trajectories in Figure 5-6 an overshoot can be seen when converging in the x and y dimension. At the end of the trajectory three and four in

Figure 5-6 there is also an overshoot in the speed control of θ. That is the reason Cody seems to orientate himself past the orientation for the reference position and reference orientation.

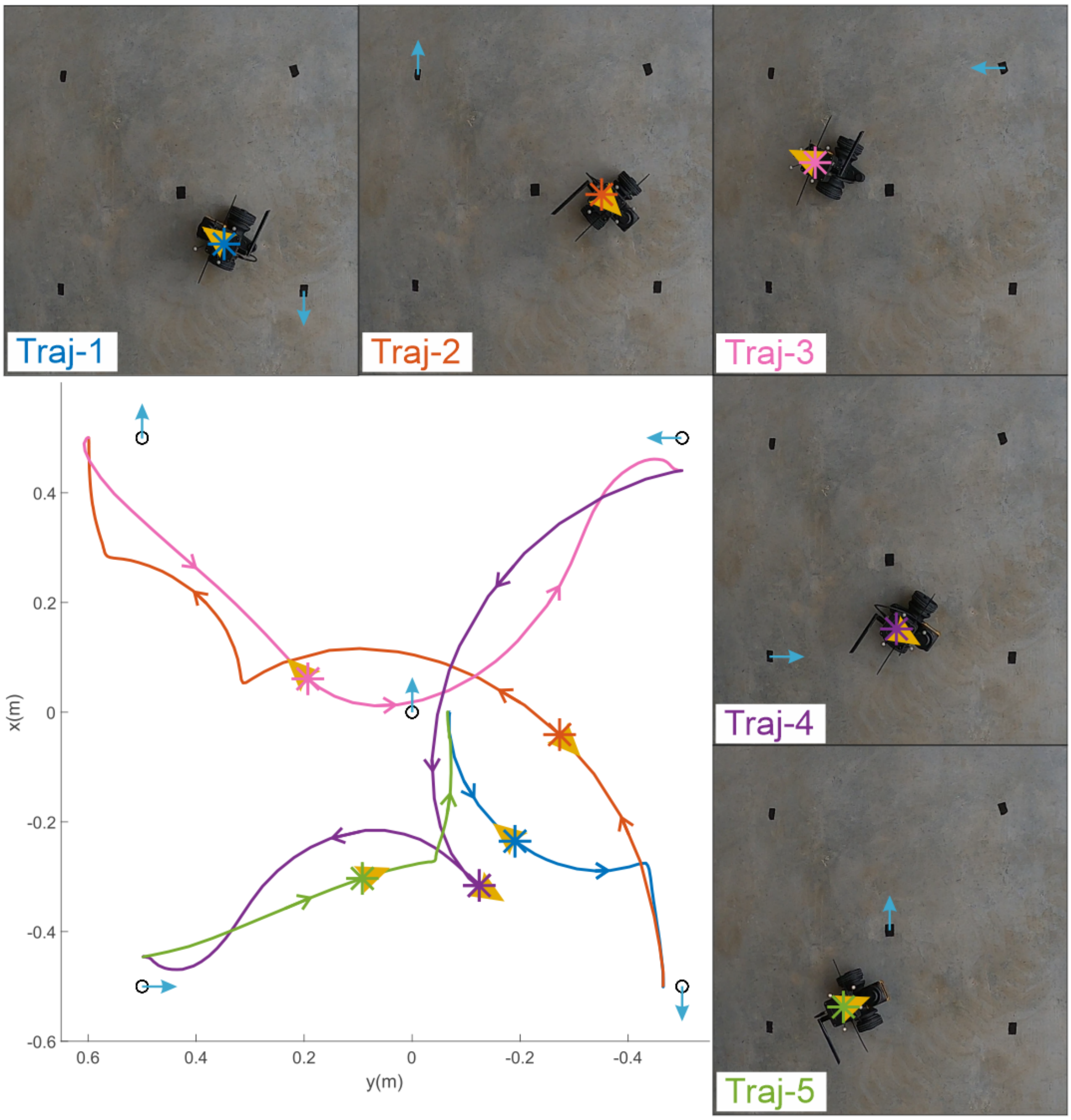


Figure 5-6: The big figure contains one sample of each trajectory. The marks indicated by ($*$) correspond to the position in the picture of Cody. The arrows in the trajectories indicate the movement of the robot along the trajectory. The large yellow arrow on the asterisk indicates the orientation of Cody. The blue arrows indicate a reference point with orientation. The blue arrow in a photo is the reference point Cody is being controlled to at that point.

In Figure 5-7, a value of 0.005 for ϵ is used for HAIC. Unfortunately, not all 50 samples were able to converge the mobile robot to the desired location. For trajectory 3 and 4, the controller was unable to converge the system to the desired state, as can be seen from the relative high error shown in the left graph of Figure 5-7. The issue with trajectory 3, was the inability to converge the θ state to zero. The issue with trajectory 4, is an issue relating to $\mu_{r,\theta}$. Further elaboration on these two issues is done in the following subsections.

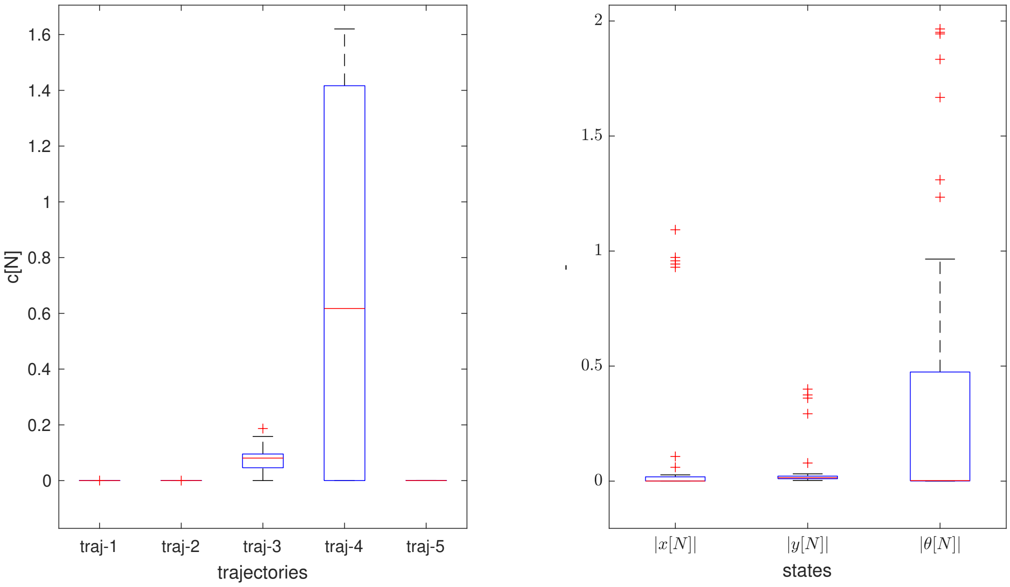


Figure 5-7: Data samples of the converged mobile robot for the value of $\epsilon = 0.005$. (Left) Five box plots of data are shown corresponding to the five trajectories elaborated in section 5-1, each box plot contains 10 data samples. (Right) Three box plots containing data corresponding to the three position states of the robot at the end of the control loop, each box plot contains 50 data samples.

A conclusion about the trade-off made with the selection of ϵ can be made when comparing the successful samples of $\epsilon = 0.005$ with the samples of $\epsilon = 0.05$. In Figure 5-8, data of only the successful controlled samples are used. Note that here, the box plots per trajectory span a much smaller as compared to Figure 5-5. As well as the box plot for $|y[N]|$ which span up to a value of 0.03 in Figure 5-8 compared to that of a value of 0.12 in Figure 5-5. Note that $|x[N]|$ and $|\theta[N]|$ have much similar range of values when comparing them for Figure 5-5 and Figure 5-8. Confirming that indeed ϵ gives a trade-off between robustness and allowable error in the y dimension of the reference frame.

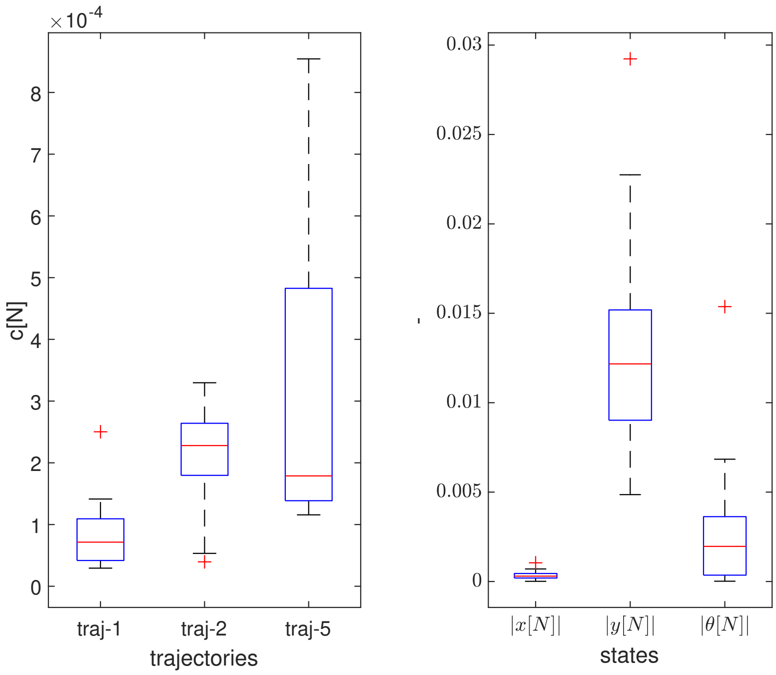


Figure 5-8: Data samples of the converged mobile robot for the value of $\epsilon = 0.005$. (Left) Three box plots of data are shown corresponding to the three successful trajectories, each box plot contains 10 data samples. (Right) Three box plots containing data corresponding to the three position states of the successful control samples, each box plot therefore contains 30 data samples.

Failure of convergence due to ϵ.

When checking the values for θ, $\mu_{r,\theta}$, and $\sqrt{(\mu_x')^2 + (\mu_y')^2}$ a conclusion can be drawn whether a similar situation happens as with the simulations in section 4-2. These beliefs and states are plotted for all ten samples obtained using HAIC with $\epsilon = 0.005$ for trajectory 3 in Figure 5-9. From all samples only one converged θ to zero. The only successful sample trajectory is indicated with a thick red line. As can be seen in the second graph, it is the only trajectory for which $\mu_{r,\theta}$ obtains a constant value of zero. Which is caused by the value of $\sqrt{(\mu_x')^2 + (\mu_y')^2}$ diving under the value of ϵ and staying there.

This proves that the conclusion obtained in chapter 4 is also applicable for a physical system. The fact that the system can converge for traj-1, traj-2, and traj-5 also indicates that the success of convergence is highly dependent on the starting position. This can indicate that, although not present in the samples obtained for $\epsilon = 0.05$, there can be a starting position which can cause the system to not converge to the origin.

To prevent the need to check for an infinite number of possible starting positions, future work can be done in the direction of theoretically proving possible convergence for a certain value of ϵ. Allowing ϵ to be chosen based on a proof instead of manual tuning would increase the robustness of this controller and allow for easier application in various systems.

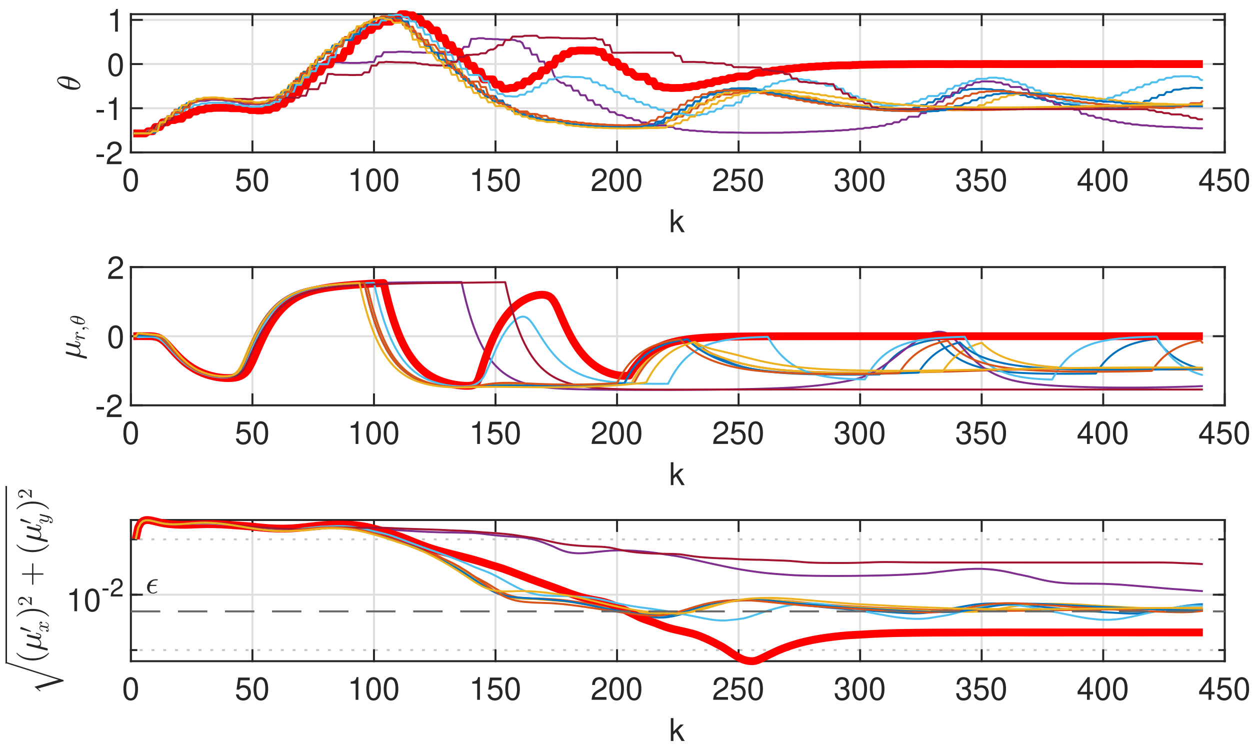


Figure 5-9: Ten samples for trajectory three where HAIC has a value of $\epsilon = 0.005$. Each colour indicates one sample trajectory, lines with the same colour in different graphs correspond to the same sample. (Top) The evolution of θ is depicted for which one sample (the thick red line) converges to zero. (Middle) The evolution of $\mu_{r,\theta}$ is depicted of which only one sample (the thick red line) converges to zero. (Bottom) The evolution of $\sqrt{(\mu_x')^2 + (\mu_y')^2}$ is depicted of which only one sample (the thick red line) obtains a consistent value lower than ϵ.

Failure of convergence due to $\mu_{r,theta}$.

In Figure 5-10, the ten samples with their values for x, v, and ω are shown. Five of the ten sample trajectories do not converge to the origin. This can be seen by checking that only five graphs reduce to 0 in the x-dimension, while all the control inputs converge to 0 and stay there. The main reason for this is the update for $\mu_{r,\theta}$ and the inability of the system to not converge the θ dimension in the previous trajectory.

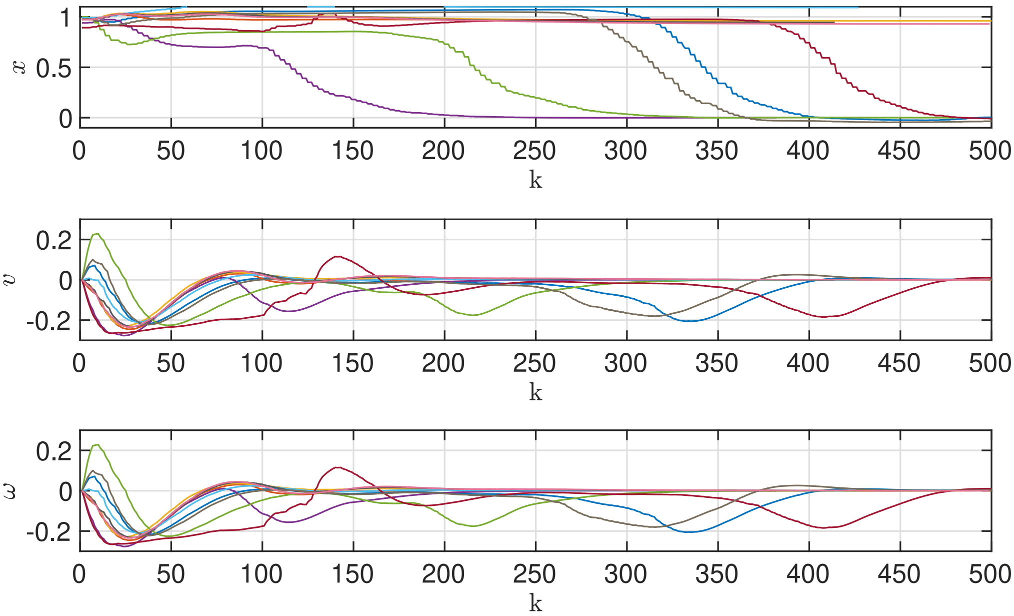


Figure 5-10: Ten sample trajectories corresponding to trajectory 4, each colour is connected to one sample. (Top) shows the evolution of x for ten different samples, only five of the ten samples were able to converge x to zero. (Middle) The evolution of the control input v is shown, which all converge to zero. (Bottom) Similar to v, the control input ω is shown, which all also converge to zero. Both the control inputs converging to zero indicate the system is converged and to set point and will not move.

In Figure 5-11, the values for the dimension x and θ of one of the five non-convergent sample trajectories are shown. In this figure everything is converged, but $\mu_{r,\theta}$ is not tracking o_T properly ($o_T = arctan(\mu'_y/\mu'_x)$). This error is caused by the update rule of $\mu_{r,\theta}$ and the inability to converge in previous trajectory. The update rule for $\mu_{r,\theta}$ is dependent on $\mu'_{r,\theta}$ as is shown again in equation (5-2).

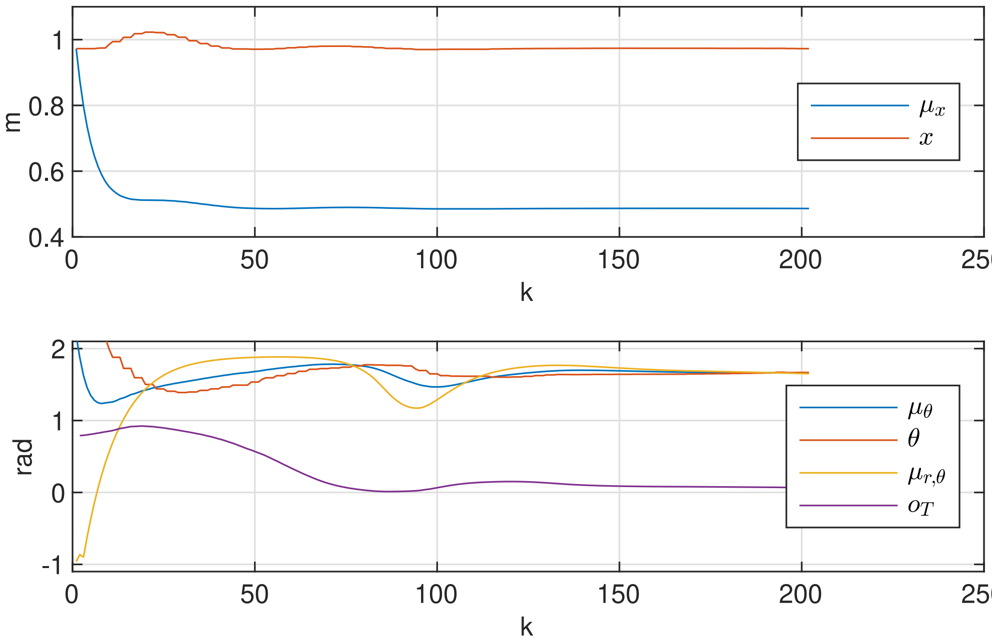


Figure 5-11: One sample trajectory which failed to converge to the origin for trajectory 4 is shown, (Top) The value for x fails to converge to zero even though the value for μ_x indicates the controller lowers the value. (Bottom) The values for θ and μ_θ are tracking $\mu_{r,\theta}$. However $\mu_{r,\theta}$ is not tracking o_T as would be expected.

$$\begin{bmatrix}\dot{\mu}_{r,\theta}\\ \dot{\mu}'_{r,\theta}\end{bmatrix} = \begin{bmatrix}\mu'_{r,\theta}\\ 0\end{bmatrix} - \kappa_{\mu_r}\begin{bmatrix}-\sigma_{o_T}^{-1}\left(arctan\left(\frac{\mu'_y}{\mu'_x}\right) - \mu_{r,\theta}\right) + \sigma_{\mu_{r,\theta}}^{-1}(\mu_{r,\theta} + \mu'_{r,\theta})\\ \sigma_{\mu_{r,\theta}}^{-1}(\mu_{r,\theta} + \mu'_{r,\theta})\end{bmatrix} \tag{5-2}$$

HAIC failed tot converge in the θ dimension for traj-3. A new goal state is given and the control for traj-4 is initiated. Because of the non convergence to zero for θ, the values of $\mu_{r,\theta}$ and $\mu'_{r,\theta}$ are non-zero. When HAIC has a value of $\sqrt{(\mu'_x)^2 + (\mu'_y)^2} > \epsilon$ the dynamics of $\mu_{r,\theta}$ with $\kappa_{\mu_r} = 0$ and $\sigma_{o_T} = 0$ are:

$$\dot{\mu}_{r,\theta} = \mu'_{r,\theta} + \left(\underbrace{arctan\left(\frac{\mu'_y}{\mu'_x}\right)}_{o_T} - \mu_{r,\theta}\right) \tag{5-3}$$

Note again that $\mu'_{r,\theta}$ is non-zero, and therefore $\dot{\mu}_{r,\theta}$ is biased and not solely based on reducing the error $(o_T - \mu_{r,\theta})$. Thus explaining the inability of $\mu_{r,\theta}$ to track o_T as shown in Figure 5-11.

Now for the particular case of this experiment, this issue can be resolved by adapting the algorithm to correctly initialize every μ to zero at the start of the control loop when there is no corresponding observation. This would create the pseudo algorithm as in Algorithm 3.

Algorithm 3 Hierarchical Active Inference Control

Initialize:
$\mu[0] = o[0], \quad \mu'[0] = \dot{o}[0], \quad \mu''[0] = 0, \quad \mu_{r,\theta}[0] = 0, \quad \mu'_{r,\theta} = 0$
Input: $\tilde{o}[k]$
Output: $u[k+1]$
$\dot{\tilde{\mu}}[k] = \frac{d}{dt}\tilde{\mu}[k] - \kappa_\mu \frac{\partial \mathcal{F}}{\partial \mu}(\tilde{\mu}[k], \tilde{o}[k], \mu_{r,\theta}[k])$
$\tilde{\mu}[k+1] = \tilde{\mu}[k] + \dot{\tilde{\mu}}[k]T$
$\dot{\tilde{\mu}}_{r,\theta}[k] = \frac{d}{dt}\tilde{\mu}_{r,\theta}[k] - \kappa_{\mu_{r,\theta}} \frac{\partial \mathcal{F}_T}{\partial \mu}(\tilde{\mu}_{r,\theta}[k], \tilde{\mu}[k])$
$\tilde{\mu}_{r,\theta}[k+1] = \tilde{\mu}_{r,\theta}[k] + \dot{\tilde{\mu}}_{r,\theta}[k]T$
$\dot{u}[k] = -\kappa_u \frac{\partial \mathcal{F}}{\partial o}\frac{\partial o}{\partial u}(\tilde{\mu}[k], \tilde{o}[k])$
$u[k+1] = u[k] + \dot{u}[k]T$
save $\tilde{\mu}[k+1], \tilde{\mu}_{r,\theta}[k+1], u[k+1]$

The issue still arises that whenever the belief of $\mu'_{r,\theta}$ of HAIC will be non-zero, and the mobile robot may for whatever reason obtain a value of $\sqrt{(\mu'_x)^2 + (\mu'_y)^2}$ larger than ϵ, a similar error will occur. The evolution of $\mu_{r,\theta}$ will not be solely focused on reducing the error of $(o_T - \mu_{r,\theta})$. Thus, having a chance that the system will again converge to a random position inside the workspace.

This indicates that HAIC cannot always converge when disturbances are present. It also indicates that the structure used for the top-layer Active Inference Control (AIC) can be improved. Further research regarding the decision function used for σ_{o_T} and $\sigma_{\mu_{r,\theta}}$ and the evolution and dynamics given to $\mu_{r,\theta}$ could be done to make sure HAIC becomes more robust. As of now, HAIC would not be robust when any disturbances are present which will cause sudden large changes in the observations.

5-3 Conclusion

The experiment done with Cody shows that HAIC works on a physical robot. A similar behaviour as with the simulations was shown with the two values for ϵ and the change in response of the system in the y and θ dimension. Both of these observations indicate that HAIC works for the control of a two-wheeled mobile robot to three-dimensional reference states, and shows the need to tune the parameter ϵ to be able to converge the system. Where the selection of ϵ is a trade-off between robustness and allowable error in the y dimension of the reference frame.

The experiments show one of the disadvantages of HAIC in its current implementation, the dynamics $\mu'_{r,\theta}$ are not robust enough to allow convergence of the system when disturbances cause sudden value changes in the observations. Although the experiment does not apply direct disturbances to HAIC, an explanation is given that there are disturbances for which HAIC will not converge to the given reference point. This therefore answers the question: “How robust is the proposed HAIC controller to noise and sudden value changes present in the measurements?” The HAIC controller is robust to noise but not to sudden large value changes present in the measurements.

It is also this error which gives a direction for future work relating to HAIC. The current implementation of HAIC based on AIC and Hierarchical Active Inference (HAI) is not robust against the large disturbances, and therefore an improvement on how to deal with these disturbances can be investigated.

Chapter 6

Conclusion

This Master Thesis has introduced Hierarchical Active Inference Control (HAIC) as a control method for the control of a nonholonomic system. HAIC is a control technique which does not require heavy computations during online operation, and has a minimal number of hyperparameters requiring selection. This allows for an easy-to-use control method, which can incorporate the nonholonomic constraint of a system. In this thesis, it was demonstrated that Hierarchical Active Inference Control (HAIC) can control a two-wheeled differential-drive mobile robot to a three-dimensional reference state.

Active Inference Control (AIC) is unable to control a mobile robot to a three-dimensional reference state. The limitation of AIC for controlling a nonholonomic mobile to a three-dimensional reference state is not taking into account the dependency of θ for the control of the dimensions x and y. This dependency is present in the form of the nonholonomic constraint in the mobile robot.

The nonholonomic system was successfully controlled by making use of hierarchical layers of AIC using a framework based on Hierarchical Active Inference (HAI). A top layer was used to incorporate the nonholonomic constraint. The introduction of a decision variable ϵ was required. The variable allows the controller to switch from prioritizing reduction of the error in the x and y dimension to prioritizing the reduction of the error in the θ dimension. The success of the HAIC controller was shown by demonstrating control of a two-wheeled mobile robot for varying initial states, of which for all it successfully converged to a reference state when the hyperparameter ϵ was correctly selected. The control method was applied both in simulation and physical application. Thus answering the question, "How can AIC and HAI be combined such that it can successfully control a two-wheeled mobile robot to a three-dimensional reference state?"

The importance of the hyperparameter ϵ was investigated. Which confirmed that ϵ has to be set in addition to AIC, which value is influenced by noise, such that convergence can be obtained when using HAIC for the control of a two-wheeled mobile robot. Known hyperparameters of AIC influence the speed and tracking ability according to previous research[1, 28]. However, HAIC for a two-wheeled differential drive mobile robot introduces the new hyperparameter ϵ. It was shown that ϵ requires selection based on noise present in the observations used in HAIC. Where the higher the noise, the higher the value for ϵ is needed to be able to converge the mobile robot to a reference state. When ϵ was not set correctly, the system would not converge the θ dimension. Both success and failure for incorrect values were shown in simulation and physical experiments. It was also shown that the value for ϵ is a trade-off between robustness against noise and possible error in the y dimension.

In results shown by the simulation and physical experiment, robustness against noise was shown based on ϵ. However, HAIC was not robust against sudden displacement or any other form which causes the observation to suddenly obtain large changes in value. This was due to the dynamics in the top layer of HAIC. Therefore, the proposed HAIC controller is robust against noise present in the measurements, but is not robust against large sudden value changes present in the measurements.

6-1 Discussion And Future Work

In this master thesis, HAIC has been introduced and proven to converge based on its demonstration on a two-wheeled differential-drive mobile robot. The focus of this thesis has been on the development and the practical application of HAIC. However, stability and convergence of HAIC have not been investigated from a theoretical basis. Future work could therefore contain this theoretical investigation and could add to a more detailed scope of the limitations and capabilities of HAIC.

The introduction of two layers of HAIC with the decision variable ϵ complicates the control method compared to that of AIC. The variable ϵ and the effect it has on σ_{o_T} and $\sigma_{\mu_{r,\theta}}$ is new when compared to previous literature relating AIC. In section 5-2, it can be seen that the dynamics caused by ϵ, made HAIC less robust against disturbances. Future work could be to investigate other possible structures while still utilizing the main characteristic of HAIC which is combining AIC and HAI.

With the introduction of ϵ a trade-off is made between robustness against noise present in the measurements and a possible error in y dimension. The main reason for not having the ability to reduce error in the y-dimension is due to the nonholonomic constraint. By introducing a horizon over which the control is calculated, a lower error in the y dimension could be possible. The horizon introduces reassurance of a low y error in advance.

In section 4-1, it was shown that the bias present in AIC continuously evolved throughout HAIC. Research has already been done in the direction of removing this bias [28]. Perhaps researching this unbiased application for HAIC could drastically improve the performance.

This thesis has focused on elaborating on HAIC and showing its success for the control of a nonholonomic system. Although HAIC underperformed compared to Model Predictive Control (MPC), a comparison with control methods making use of a similar computation load would be interesting. This would also allow the further exploration of choosing optimal values for the hyperparameters, which could even evolve to introducing a best practice tuning method for HAIC. A possible introduction of an online update rule for some hyperparameters as already done by Baioumy et al. in previous research[28]. This also gives a possible indication for the online adaptive tuning of hyperparameters for HAIC.

The hierarchical aspect of HAIC allows the nonholonomic constraint to be considered when controlling a two-wheeled mobile robot. A similar capability can be expected when applied to different nonholonomic systems. This master thesis is also a step in the direction of applying HAIC for the use of obstacle avoidance. Where the obstacle is given as a similar constraint which the controller has to take into account. However, implementation will be slightly different. In this thesis, the constraint is upheld by the system, and not the controller. While for obstacle avoidance, it will be the controller which should uphold the constraint instead of the system.

The hierarchical aspect can be even further investigated by adding additional layers. This can be done in multiple forms, but one example would be that of multi-agent control using HAIC. Where the top layer introduces the combined goal of multiple agents and distributes the control for every singular agent to a lower layer of AIC.

Appendix A

Appendix

A-1 Simplification Of The Free Energy

The simplification of the free energy is explained according to the steps based on research done by Buckley et al. [34]

From chapter 3, the variational free energy was reduced to.

$$\mathcal{F} = \int q(z) \ln q(z) dz - \int q(z) \ln p(z, o) dz \tag{A-1}$$

Note that the normal Gaussian distribution $q(z)$ to approximate $p(z|o)$ is given as:

$$q(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \tag{A-2}$$

Also, the way to calculate the variance σ^2 of a distribution $p(x)$ with mean μ is:

$$\sigma^2 = \int (x - \mu)^2 p(x) dx \tag{A-3}$$

The first term of (A-1) can be simplified by using $\int q(z)dz = 1$ for any normal distribution and using the formula for the variance. This allows us to reduce the first term by writing out $\ln(q(z))$ using equation (A-2).

$$\begin{aligned}
\int q(z) \ln q(z) dz &= \int q(z) \left(-\frac{1}{2} \ln (\sigma^2 2\pi) - \frac{(z-\mu)^2}{2\sigma^2} \right) dz \\
&= -\frac{1}{2} \ln (\sigma^2 2\pi) \underbrace{\int q(z) dz}_{=1} - \frac{1}{2\sigma^2} \underbrace{\int q(z)(z-\mu)^2 dz}_{=\sigma^2} \\
&= -\frac{1}{2} \ln (\sigma^2 2\pi) - \frac{1}{2}
\end{aligned} \tag{A-4}$$

The second term is contain the term $\ln\,p(z,o)$ which is still unspecified. What is known, is that $q(z)$ sharply peaks at μ. By assuming that $\ln\,p(z,o)$ is a smooth function, a Taylor expansion can be used around μ because it's non-zero only at its peak. This would obtain the following equation:

$$\begin{aligned}\int q(z)\,\ln\,p(z,o)dz = \int q(z)dz\Bigg(\ln\,p(\mu,o) + \left[\frac{d}{dz}\ln\,p(z,o)\right]_\mu (z-\mu)\ ...\\ +\frac{1}{2}\left[\frac{d^2}{dz^2}\ln\,p(z,o)\right]_\mu (z-\mu)^2\Bigg)\end{aligned} \tag{A-5}$$

Again, the fact that $\int q(z)dz = 1$ and equation for the variance can be used to simplify the equation to:

$$\int q(z)\,\ln\,p(z,o)dz = \ln\,p(\mu,o) + \frac{1}{2}\left[\frac{d^2}{dz^2}\ln\,p(z,o)\right]\sigma^2 \tag{A-6}$$

Combining simplifications of the first and second term:

$$\mathcal{F} = -\ln\,p(\mu,o) - \frac{1}{2}\left(\left[\frac{d^2}{dz^2}\ln\,p(z,o)\right]_\mu \sigma^2 + \ln\,(\sigma^2 2\pi) + 1\right) \tag{A-7}$$

Note that the function of $\mathcal{F}$ is now dependent on μ, o, and σ^2. The dependency of the variance σ^2 can be removed by already obtaining the optimal variance for reducing the free energy $\mathcal{F}$. This is done by stating $\frac{\partial\mathcal{F}}{\partial\sigma^2} = 0$:

$$\begin{aligned}\frac{\partial\mathcal{F}}{\partial\sigma^2} &= \frac{1}{2}\left(\left[\frac{d^2}{dz^2}\ln\,p(z,o)\right]_\mu + \frac{1}{\sigma^2}\right) = 0\\ \sigma^2 &= -\left[\frac{d^2}{dz^2}\ln\,p(z,o)\right]_\mu^{-1}\end{aligned} \tag{A-8}$$

The free energy function $\mathcal{F}$ is now a function of μ and o (σ^2 is kept but also a function of μ and o):

$$\mathcal{F} = -\ln\,p(\mu,o) - \frac{1}{2}\ln\,(\sigma^2 2\pi) \tag{A-9}$$

The last step is to use the approximation of $-\ln\,p(\mu,o)$ to now approximate the variational free energy $\mathcal{F}$, thus reducing down to:

$$\mathcal{F} \approx -\ln\,p(\mu,o) \tag{A-10}$$

Bibliography

[1] C. Pezzato, R. Ferrari, and C. H. Corbato, "A novel adaptive controller for robot manipulators based on active inference," *IEEE Robotics and Automation Letters*, vol. 5, pp. 2973–2980, 4 2020.

[2] A. M. Research, "Inspection robots market," 2021.

[3] Research and Markets, "Worldwide inspection robots industry to 2029 - impact analysis of drivers and restraints," 2022.

[4] S. Bonadies, A. Lefcourt, and S. A. Gadsden, "A survey of unmanned ground vehicles with applications to agricultural and environmental sensing," vol. 9866, p. 98660Q, SPIE, 5 2016.

[5] V. A. Jorge, R. Granada, R. G. Maidana, D. A. Jurak, G. Heck, A. P. Negreiros, D. H. dos Santos, L. M. Gonçalves, and A. M. Amory, "A survey on unmanned surface vehicles for disaster robotics: Main challenges and directions," *Sensors (Switzerland)*, vol. 19, 2 2018.

[6] X. Zhou, X. Yu, Y. Zhang, Y. Luo, and X. Peng, "Trajectory planning and tracking strategy applied to an unmanned ground vehicle in the presence of obstacles," *IEEE Transactions on Automation Science and Engineering*, vol. 18, pp. 1575–1589, 10 2021.

[7] N. Muchiri and S. Kimathi, "A review of applications and potential applications of uav," pp. 280–283, 2016.

[8] C. Torresan, A. Berton, F. Carotenuto, S. F. D. Gennaro, B. Gioli, A. Matese, F. Miglietta, C. Vagnoli, A. Zaldei, and L. Wallace, "Forestry applications of uavs in europe: a review," *International Journal of Remote Sensing*, vol. 38, pp. 2427–2447, 5 2017.

[9] A. Thibbotuwawa, P. Nielsen, B. Zbigniew, and G. Bocewicz, "Energy consumption in unmanned aerial vehicles: A review of energy consumption models and their relation to the uav routing," vol. 853, pp. 173–184, Springer Verlag, 2019.

[10] P. Beigi, M. S. Rajabi, and S. Aghakhani, “An overview of drone energy consumption factors and models,” *arXiv:2206.10775v2*, 6 2022.

[11] M. F. Silva and J. A. MacHado, “A literature review on the optimization of legged robots,” *Journal of Vibration and Control*, vol. 18, pp. 1753–1767, 10 2012.

[12] G. Fadini, T. Flayols, A. D. Prete, N. Mansard, and P. Souères, “Computational design of energy-efficient legged robots: Optimizing for size and actuators,” vol. 2021-May, pp. 9898–9904, Institute of Electrical and Electronics Engineers Inc., 2021.

[13] C. Samson and K. Ait-Abderrahim, “Mobile robot control. part 1: Feedback control of nonholonomic wheeled cart in cartesian space,” 1990.

[14] R. J. Stern, “Brockett’s stabilization condition under state constraints,” 2002.

[15] J. M. Yang and J. H. Kim, “Sliding mode motion control of nonholonomic mobile robots,” *IEEE Control Systems*, vol. 19, pp. 15–23, 1999.

[16] A. Azzabi and K. Nouri, “Design of a robust tracking controller for a nonholonomic mobile robot based on sliding mode with adaptive gain,” *International Journal of Advanced Robotic Systems*, vol. 18, 2021.

[17] T. Fukao, H. Nakagawa, and N. Adachi, “Adaptive tracking control of a nonholonomic mobile robot,” *IEEE Transactions on Robotics and Automation*, vol. 16, pp. 609–615, 10 2000.

[18] T. C. Lee, K. T. Song, C. H. Lee, and C. C. Teng, “Tracking control of unicycle-modeled mobile robots using a saturation feedback controller,” *IEEE Transactions on Control Systems Technology*, vol. 9, pp. 305–318, 3 2001.

[19] Y. Wang, Z. Miao, H. Zhong, and Q. Pan, “Simultaneous stabilization and tracking of nonholonomic mobile robots: A lyapunov-based approach,” *IEEE Transactions on Control Systems Technology*, vol. 23, pp. 1440–1450, 7 2015.

[20] R. Zheng, Y. Liu, and D. Sun, “Enclosing a target by nonholonomic mobile robots with bearing-only measurements,” *Automatica*, vol. 53, pp. 400–407, 3 2015.

[21] M. Michałek and K. Kozłowski, “Vector-field-orientation feedback control method for a differentially driven vehicle,” *IEEE Transactions on Control Systems Technology*, vol. 18, pp. 45–65, 1 2010.

[22] T. P. Nascimento, C. E. Dórea, and L. M. G. Gonçalves, “Nonholonomic mobile robots’ trajectory tracking model predictive control: A survey,” *Robotica*, vol. 36, pp. 676–696, 5 2018.

[23] K. Worthmann, M. W. Mehrez, M. Zanon, G. K. Mann, R. G. Gosine, and M. Diehl, “Model predictive control of nonholonomic mobile robots without stabilizing constraints and costs,” *IEEE Transactions on Control Systems Technology*, vol. 24, pp. 1394–1406, 7 2016.

[24] K. Friston, T. FitzGerald, F. Rigoli, P. Schwartenbeck, J. O’Doherty, and G. Pezzulo, “Active inference and learning,” *Neuroscience and Biobehavioral Reviews*, vol. 68, pp. 862–879, 9 2016.

[25] G. Pezzulo, F. Rigoli, and K. J. Friston, "Hierarchical active inference: A theory of motivated control," *Trends in Cognitive Sciences*, vol. 22, pp. 294–306, 4 2018.

[26] L. Pio-Lopez, A. Nizard, K. Friston, and G. Pezzulo, "Active inference and robot control: A case study," *Journal of the Royal Society Interface*, vol. 13, 9 2016.

[27] P. Lanillos, C. Meo, C. Pezzato, A. A. Meera, M. Baioumy, W. Ohata, A. Tschantz, B. Millidge, M. Wisse, C. L. Buckley, and J. Tani, "Active inference in robotics and artificial agents: Survey and challenges," 12 2021.

[28] M. Baioumy, P. Duckworth, B. Lacerda, and N. Hawes, "Active inference for integrated state-estimation, control, and learning," vol. 2021-May, pp. 4665–4671, Institute of Electrical and Electronics Engineers Inc., 2021.

[29] M. Baioumy, C. Pezzato, R. Ferrari, and N. Hawes, "Unbiased active inference for classical control," 2022.

[30] T. Matsumoto, W. Ohata, F. C. Benureau, and J. Tani, "Goal-directed planning and goal understanding by extended active inference: Evaluation through simulated and physical robot experiments," *Entropy*, vol. 24, 4 2022.

[31] O. Çatal, T. Verbelen, T. V. de Maele, B. Dhoedt, and A. Safron, "Robot navigation as hierarchical active inference," *Neural Networks*, vol. 142, pp. 192–204, 10 2021.

[32] R. W. Brockett, "Asymptotic stability and feedback stabilization," *Differential geometric control theory*, vol. 27(1), pp. 181–191, 1983.

[33] M. D. Kvalheim and D. E. Koditschek, "Necessary conditions for feedback stabilization and safety," 6 2022.

[34] C. L. Buckley, C. S. Kim, S. McGregor, and A. K. Seth, "The free energy principle for action and perception: A mathematical review," *Journal of Mathematical Psychology*, vol. 81, pp. 55–79, 12 2017.

[35] Avular, "Avular - mobile robotics," 2023.

[36] E. Langford, "Quartiles in elementary statistics," *Journal of Statistics Education*, vol. 14, 2006.

Glossary

List of Acronyms

AIC	Active Inference Control
HAI	Hierarchical Active Inference
HAIC	Hierarchical Active Inference Control
NN	Neural Network
PID	Proportional Integral Derivative
KL	Kullback-Leibler
WMR	wheeled mobile robot
SMC	sliding mode control
MPC	Model Predictive Control
UAVs	unmanned air vehicles
UGVs	unmanned ground vehicles
IQR	Interquartile Range
LTI	linear time-invariant