University of Nevada, Reno

DISTRIBUTED CONTROL OF MULTI-AGENT SYSTEMS USING BIOLOGICALLY-INSPIRED REINFORCEMENT LEARNING

A Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of Doctor of Philosophy in

Electrical Engineering

by

Mohammad Jafari

Dr. Hao Xu / Dissertation Advisor

August 2018

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

MOHAMMAD JAFARI

Entitled

Distributed Control Of Multi-Agent Systems Using Biologically-Inspired Reinforcement Learning

be accepted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

Hao Xu, Ph.D., Advisor

Luis Rodolfo GARCIA CARRILLO, Ph.D., Committee Member

Sesh Commuri, Ph.D., Committee Member

Yantao Shen, Ph.D., Committee Member

Monica Nicolescu, Ph.D., Graduate School Representative

David W. Zeh, Ph. D., Dean, Graduate School
August, 2018

ABSTRACT

In this dissertation, we investigate the real-time flocking control of Multi-Agent Systems (MAS) in presence of system uncertainties and dynamic environment. To handle the impacts from system uncertainties and dynamic environment, a novel reinforcement learning technique, which is appropriate for real-time implementation, has been integrated with multi-agent flocking control in this dissertation. The Brain Emotional Learning Based Intelligent Controller (BELBIC) is a biologically-inspired reinforcement learning based controller relying on a computational model of emotional learning in mammalian limbic system. The learning capabilities, multi-objective properties, and low computational complexity of BELBIC make it a very promising learning technique for implementation in real-time applications. Firstly, a novel brain emotional learning based flocking control structure is proposed. Then, the real-time update laws are developed to tune the emotional signals based on the real-time operation data. It is important to note that this data-driven reinforcement learning approach relaxes the requirement for system dynamics and effectively handle the uncertain impacts of environment. Using the tuned emotional signals, the optimal flocking control can be obtained. The Lyapunov analysis has been used to prove the convergence of the proposed design. The effectiveness of the proposed design is also demonstrated through numerical and experimental results based on the coordination of multiple Unmanned Aircraft Systems (UAS) platforms.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my advisor Dr. Hao Xu, for inspiring and encouraging me to continue and complete this work. Thank you for all your support throughout this dissertation work especially during hard times. I can never thank him enough for giving me the freedom to work on the fields of my interest and helping me to pursue my goals. I am sure your advices will help me in my personal and professional life in the long run.

I wish to give my very special thanks to Dr. Luis Rodolfo GARCIA CARRILLO for supporting me during my study in EE Department and for all the inspiration and encouragement.

I also would like to thank committee members Dr. Monica Nicolescu, Dr. Sesh Commuri and Dr. Yantao Shen for their valuable inputs in improving the dissertation work.

I would also like to thank Electrical and Biomedical Engineering department for providing all technical and material support throughout my graduate program.

I am incredibly grateful to my family. My parents and my sister gave me their unconditional love and support during all these years. While I never had the opportunity to see my family during my studies at UNR, my heart has always been with them. They never stopped believing in me and did everything they could to help me achieve my goals.

I would also like to acknowledge all my friends and colleagues for their support.

Table of Contents

List of Tables

List of Figures

"If you want to be incrementally better: Be competitive.

If you want to be exponentially better: Be cooperative.

- Anonymous"

Chapter 1

Introduction

During the past decade, diverse research communities have developed several advanced control strategies for coordination of Multi-Agent Systems (MAS), see for example [1, 3–14] and the references therein. In most of these MAS control methods, *flocking* problem have been formulated and investigated thoroughly. Flocking is the collective motion of a large number of self-propelled entities exhibited by many living beings such as birds, fish, bacteria, and insects [15]. Flocking is also considered as an emergent behavior, which is caused by a number of simple rules followed by each agent, and that does not require any central coordination.

The seminal work in [7] introduced three basic rules for simulating the flocking behavior, specifically, *separation*, *alignment*, and *cohesion*. In recent years, there has been a surge of interest among researchers to improve the flocking behavior of MAS. One can mention, for example, adaptive flocking control approaches for dealing with varying and noisy environments [16], [17], robust flocking controllers to handle model uncertainty [18, 19], and flocking control methods with the capability of disturbance handling [20].

Several critical aspects should be considered in the real-time implementation of control strategies on mobile robotic platforms [21–26], for example, model uncertainty, disturbances, energy expenses and actuator saturation. Diverse research efforts have been proposed aiming at addressing the issues arising from these practical and harsh conditions. Considering model uncertainty and disturbances, distributed tracking and estimation of MAS with uncertain dynamics has been presented in [27]. In [28], the problem of robust consensus tracking for MAS with disturbances and unmodeled dynamics has been studied. In addition, neural adaptive flocking of MAS has been addressed in [29–31]. More recently, authors in [32] investigated the application of Q-learning in leader-follower based flocking with small fixed-wing UAVs.

Also, to tackle the energy expenses and actuator saturation related problems, the authors in [33] presented a decentralized approach to perform formation maneuvers by groups of mobile robots taking into account the actuator saturation. Closely related, a flocking control with constraints is proposed in [34]. Swarm aggregation of MAS with actuator saturation has been addressed in [35,36]. Furthermore, a leader-following tracking problem for MAS with a varying-velocity leader and input saturation was investigated in [37]. In [38], the authors introduced a decentralized connectivity maintenance strategy for mobile networks with constrained inputs. Recently, an energy function-based approach for estimating the required control force for network connectivity preservation and collision avoidance was presented in [39].

In general, flocking strategies available in the literature are addressing, among others i.e., (i) the optimization of the MAS flocking control, (ii) the robustness for dealing with the dynamic environment, and (iii) the capability of dealing with

MAS uncertainty. However, most of these recent approaches are not designed for having multi-objective properties i.e., they only considered solving one of the aforementioned problems. Indeed, when multiple goals are targeted in parallel to the flocking problem, the computational complexity of the overall problem is not suitable for real-time implementation. Additionally, most of the existing works need the knowledge of system dynamics and not applicable in presence of dynamic environment.

Therefore, to overcome this deficiency, biologically-inspired methods have been increasingly employed to solve complex computational problems. Brain Emotional Learning Based Intelligent Controller (BELBIC) is one of the most promising approaches that adopts the learning model developed in [40] in order to mimic the functionalities of the brain that are known to produce emotion, i.e., the *amygdala*, *orbitofrontal cortex*, *thalamus*, and *sensory input cortex*. Strategies based on BELBIC have shown to be a very promising solution in terms of robustness and uncertainty handling [41–43] due to the flexibility of its architecture. BELBIC has two main inputs: *Sensory Inputs* (SI) and *Emotional Signal* (ES). Properly adjusting both SI and ES, makes this controller an appealing strategy for addressing real-time applications [41]. Furthermore, BELBIC architecture has the computational complexity on the order of $O(n)$ [44] that is much smaller and better for real-time implementation compared with other existing learning-based intelligent control methods.

1.1 Contributions

This work contributes to the research of multi-agent systems by developing intelligent cooperative control in a distributed fashion to realize coordinated motion

control. This dissertation is presented in five papers, and their contributions is provided below. The common theme in the five papers is the Brain Emotional Learning (BEL) model.

1.1.1 A Neurobiologically-inspired Intelligent Trajectory Tracking Control for Unmanned Aircraft Systems with Uncertain System Dynamics and Disturbance

The main contribution of this paper consists on advancing the state of the art of the UAS intelligent control strategies which are capable of handling the system uncertainties and disturbances, and also are appropriate for real-time implementation due to its low computational complexity. In order to accomplish this goal, we take advantage of the computational model of emotional learning encountered in the mammalian limbic system, i.e., BELBIC [43, 45–50].

1.1.2 A Biologically-Inspired Reinforcement Learning based Intelligent Distributed Flocking Control for Multi-Agent Systems in Presence of Uncertain System and Dynamic Environment

The main contribution of this paper is to utilize the computational model of emotional learning in the mammalian limbic system, i.e., BELBIC, for developing and experimental validation of a novel intelligent flocking control method for practical

MAS in presence of uncertainty system dynamics and dynamic environment [51–55].

1.1.3 A Biologically-Inspired Distributed Fault Tolerant Flocking Control for Multi-Agent System in Presence of Uncertain Dynamics and Unknown Disturbance

The main contribution of this paper is to develop Resilient BELBIC (R-BELBIC) as an intelligent distributed resilient controller for flocking of MAS in presence of system uncertainties and unknown disturbances. In order to guarantee resilience, we employed the computational model of emotional learning in the mammalian limbic system, i.e., BELBIC with context [56,57].

1.1.4 A Game Theoretic Based Biologically-Inspired Distributed Intelligent Flocking Control for Multi-UAV Systems with Network Imperfections

The main contribution of this paper is to develop a model-free distributed intelligent control methodology to overcome the challenges including the unknown disturbances and uncertainties from environment and system in networked multi-UAV systems. To this end, we propose a game theoretic based biologically-inspired distributed intelligent controller, which takes advantage of the game theory and the computational model of emotional learning in the mammalian limbic system [58–

60].

1.1.5 Brain Emotional Learning-Based Path Planning and Intelligent Control Co-Design for Unmanned Aerial Vehicle in Presence of System Uncertainties and Dynamic Environment

The contribution of this paper has three main components, specifically (i) a path planning and control co-design, (ii) a learning based approach that can effectively handle the uncertainties from unstable UAV system and complex environment, and (iii) a low computational learning technique that can be implemented in real-time. To this end, we utilize the computational model of emotional learning in mammal's brain, i.e., BEL, for developing a novel path planning and intelligent control co-design for practical real-time implementation in UAVs in the presence of system uncertainties and dynamic environments [61].

Chapter 2

A Neurobiologically-inspired Intelligent Trajectory Tracking Control for Unmanned Aircraft Systems with Uncertain System Dynamics and Disturbance

In this paper, a novel neurobiologically inspired intelligent tracking controller is developed and implemented for Unmanned Aircraft Systems (UAS) in presence of uncertain system dynamics and disturbance. The methodology adopted, known as Brain Emotional Learning Based Intelligent Controller (BELBIC), is based on a novel computational model of emotional learning in mammals' brain limbic sys-

tem. Compared with conventional model-based control methods, BELBIC is more suitable for practical UAS since it can maintain the real-time UAS performance without known system dynamic and disturbance. Furthermore, the learning capability and low computational complexity of BELBIC make it very promising for implementation in complex real-time applications. Moreover, we proved that our proposed methodology guarantees the convergence. To evaluate the practical performance of proposed design, BELBIC has been implemented into a benchmark UAS. Numerical and experimental results demonstrated the applicability and satisfactory performance of the proposed BELBIC-inspired design.

2.1 Introduction

Due to recent advances in technology and the formalization of clearer regulations, Unmanned Aircraft systems (UAS) are being proposed for safely executing different kind of applications considered to be too risky for humans [11]. One of the most versatile UAS platforms is the quad rotorcraft, a vehicle with Vertical Takeoff and Landing (VTOL) and hovering capabilities [62]. Due to their popularity, UAS attracted enormous interests from diverse research communities [17, 63–65]. For example, the authors in [66] proposed an adaptive fuzzy controller for stabilizing a quad rotorcraft during flight. In addition, adaptive neural control with extreme learning machine has been proposed in [67]. Closely related, in [68] an adaptive sliding mode control approach has been utilized for controlling the flight of UAS. A reinforcement learning method has been employed by authors in [69] towards the same goal. The authors in [70] proposed a nonlinear robust output feedback tracking control for stabilizing a UAS by using a quaternion representation. Most recently, a robust attitude control approach for controlling a miniature quad rotor-

craft has been proposed by the authors in [71]. Authors in [72] employed a linear dual disturbance observer based control strategy to improve the trajectory tracking precision for an indoor quadrotor. Furthermore, an $\mathcal{L}_1$ adaptive fault tolerant controller is developed for the trajectory tracking of a quad rotorcraft in [73]. Despite the effectiveness shown by these previous methodologies, most of these works require detailed information concerning the system, i.e., the UAS dynamics, in order to be successfully implemented.

It is well known that, if the model of the system to be controlled is complex and unknown, or even if some aspects of the model are uncertain, the development of an appropriate control strategy becomes a very challenging task. Furthermore, when dealing with real-time systems, in addition to model uncertainties diverse issues such as environmental disturbances might arise. For this reason, the development of control strategies with less dependency on the full knowledge of the system dynamics is critical for successfully accomplishing the robot's mission.

Related Work

In recent years, biologically-inspired approaches have been extensively utilized as a promising solution for computationally complex problems. One of such approaches, which is based on a computational model that mimics those parts of the brain which are known to produce emotion, was introduced in [40]. The methodology, known as Brain Emotional Learning Based Intelligent Controller (BELBIC), imitates the *emotional* parts of the mammalian brain, i.e., the *amygdala*, *orbitofrontal cortex*, *thalamus*, and *sensory input cortex*. From a control systems point of view, the main interest in implementing emotional learning is due to the fact that, this kind

of methodology is effective when the model dynamics (equations of the system) are fully or partially unknown.

In [41], authors have demonstrated that BELBIC is a model-free controller which has shown promising performance when dealing with noise and system uncertainty. BELBIC has two main inputs: *Sensory Inputs* (SI) and *Emotional Signal* (ES), and the flexibility in defining both SI and ES to solve multi-objective problems [42,46,74,75] makes this controller a practical tool for implementation in real-time applications. Therefore, BELBIC could be effective for controlling a system even when the only available information for the designer are the states of system and the controller performance feedback that is available in the form of a win or lose signal [76]. Furthermore, BELBIC has a single layered architecture, and therefore, its computational complexity is on the order of $O(n)$ that is much smaller and better for real-time implementation compared with other existing learning-based intelligent control.

Main Contributions

The main contribution of this paper consists on advancing the state of the art of the UAS intelligent control strategies which are capable of handling the system uncertainties and disturbances, and also are appropriate for real-time implementation due to its low computational complexity. In order to accomplish this goal, we take advantage of the computational model of emotional learning encountered in the mammalian limbic system, i.e., BELBIC. Our recent works in [43] introduced the theoretical framework for attitude control of a quad rotorcraft using BELBIC. These results are improved and extended in this paper, by considering the problem

of stabilizing the full 6 degrees of freedom (DoF) of the UAS and also intelligent tracking control of UAS, when subjected to unknown dynamics and impacts from harsh environment, e.g., disturbances. The learning capability and low computational complexity of the proposed approach enabled its real-time implementation that has been validated in a UAS platform. The design and experimental validation presented in this paper are an original work, which have demonstrated very promising results. Ultimately, we proved that our proposed methodology guarantees the convergence.

The rest of the paper is organized as follows. Section 2.2 presents the problem formulation, as well as basic information concerning the quad rotorcraft modeling, and preliminaries about BELBIC. Our main contribution is introduced in Section 2.3, which consists of a real-time quad rotorcraft control application based on BELBIC. Sections 2.4 and 2.5 present numerical simulation results and experimental results, respectively. The conclusion of the paper and future directions of our work are provided in Section 2.6.

2.2 Problem Formulation and Preliminaries

Consider the problem of stabilizing the six DoF of a UAS, while at the same time performing a trajectory tracking control during a real-time autonomous flight. Consider also that the dynamics of the UAS are uncertain and not fully available, and that the platform is evolving in an uncertain environment experiencing external disturbances. The main objective of this research consists of the design and real-time implementation of an intelligent control strategy, which is effective for stabilizing the UAS during autonomous flights, even under the challenging cir-

cumstances previously described.

The solution proposed in this paper corresponds to a BELBIC-inspired intelligent control design, which provides the following benefits:

- a model-free controller where the knowledge of system dynamics is not required.
- uncertainty handling and noise/disturbance rejection are achievable.
- a low computational algorithm that is appropriate for real-time implementation.

The details concerning the proposed methodology are explained in the following subsections.

2.2.1 Representation of the UAS dynamics

For analytically validating the proposed controller, the dynamics of the UAS are described with respect to a hybrid system of coordinates, which we will call the H-frame [77]. This system is composed of a set of nonlinear equations expressed with respect to (w.r.t.) an inertial frame (I-frame), and a set of angular equations expressed w.r.t. a body fixed frame (B-frame), see Figure 2.1. This new reference is adopted because its easy to express the UAS dynamics in combination with the control, in particular for the vertical position, in the inertial frame. Then, the UAS

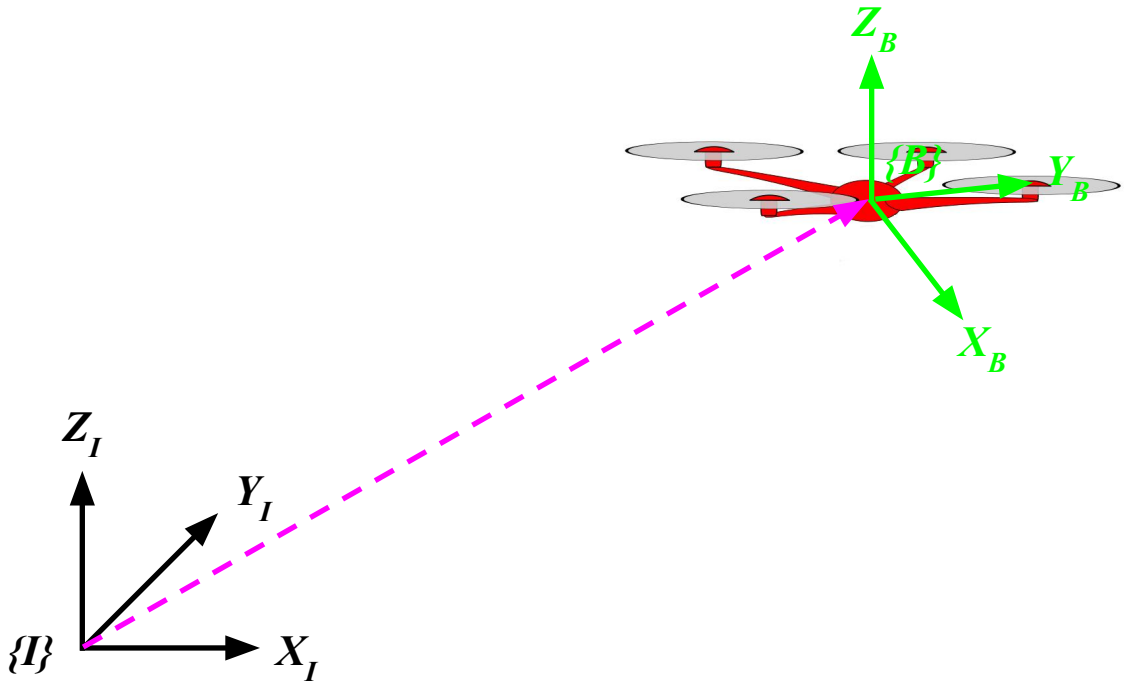

Figure 2.1: System Coordinates.

equations expressed in the H-frame are described as follows [77]:

$$
\begin{cases}
\ddot{X} = (\sin\psi\sin\phi + \cos\psi\sin\theta\cos\phi)\frac{U_1}{m} \\
\ddot{Y} = (-\cos\psi\sin\phi + \sin\psi\sin\theta\cos\phi)\frac{U_1}{m} \\
\ddot{Z} = -g + (\cos\theta\cos\phi)\frac{U_1}{m} \\
\dot{p} = \frac{I_{YY}-I_{ZZ}}{I_{XX}}qr - \frac{J_{TP}}{I_{XX}}q\mathbf{\Omega} + \frac{U_2}{I_{XX}} \\
\dot{q} = \frac{I_{ZZ}-I_{XX}}{I_{YY}}pr - \frac{J_{TP}}{I_{YY}}p\mathbf{\Omega} + \frac{U_3}{I_{YY}} \\
\dot{r} = \frac{I_{XX}-I_{YY}}{I_{ZZ}}pq + \frac{U_4}{I_{ZZ}}
\end{cases}
\tag{2.1}
$$

where θ, ϕ, and ψ are Pitch, Roll, and Yaw angles, respectively, U_1 is the collective throttle, and U_2, U_3, and U_4 are the Roll, Pitch, and Yaw moments. I_{XX}, I_{YY}, and I_{ZZ} are the body moment of inertia around the X, Y, and Z axis, respectively. g is the acceleration due to gravity, m is the UAS mass, and J_{TP} is the total rotational moment of inertia around the propeller axis. The relationship between the propellers' speed and the moments are defined by

$$U_1 = b_q(\mathbf{\Omega}_1^2 + \mathbf{\Omega}_2^2 + \mathbf{\Omega}_3^2 + \mathbf{\Omega}_4^2) \tag{2.2}$$

$$U_2 = b_q l(-\mathbf{\Omega}_2^2 - \mathbf{\Omega}_3^2 + \mathbf{\Omega}_1^2 + \mathbf{\Omega}_4^2) \tag{2.3}$$

$$U_3 = b_q l(-\mathbf{\Omega}_1^2 - \mathbf{\Omega}_2^2 + \mathbf{\Omega}_3^2 + \mathbf{\Omega}_4^2) \tag{2.4}$$

$$U_4 = d_q(-\mathbf{\Omega}_1^2 + \mathbf{\Omega}_2^2 - \mathbf{\Omega}_3^2 + \mathbf{\Omega}_4^2) \tag{2.5}$$

$$\mathbf{\Omega} = -\mathbf{\Omega}_1 + \mathbf{\Omega}_2 - \mathbf{\Omega}_3 + \mathbf{\Omega}_4 \tag{2.6}$$

where $\mathbf{\Omega}_1$, $\mathbf{\Omega}_2$, $\mathbf{\Omega}_3$, and $\mathbf{\Omega}_4$ are front-left, front-right, rear-right, and rear-left propeller speeds, respectively. Also, b_q is a thrust factor specific for the quad rotorcraft platforms, d_q is drag factor of the quad rotorcraft and l is the distance between the center of the quad rotorcraft and the center of the propeller. The nonlinearities of the quad rotorcraft motors, which is basically the relationship between the motors' voltage and the propellers' speed is described as follows

$$J_{TP}\dot{\mathbf{\Omega}} = -\frac{K_E K_M}{R}\eta N^2 \mathbf{\Omega} - d_q \mathbf{\Omega}^2 + \frac{K_M}{R}\eta N \nu \tag{2.7}$$

where K_E and K_M are the electric and mechanic motor constants, respectively, R is the motor resistance, and ν is the motor voltage. The term N, which represents a gearbox reduction ratio, and η, which is the conversion efficiency of the gearbox, may or may not be included, according to the UAS platform adopted for the real-time experiments.

2.2.2 Brain Emotional Learning-Based Intelligent Controller

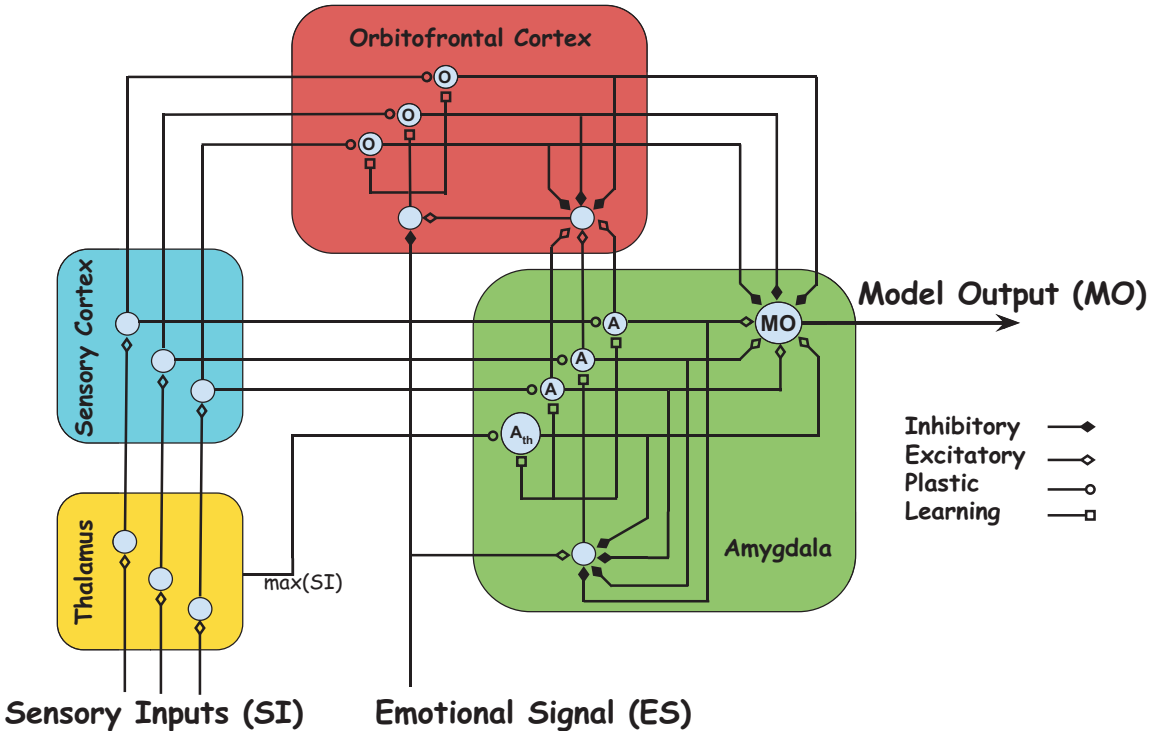

Figure 2.2: Computational model of emotional learning.

BELBIC is a neurobiologically-motivated methodology, based on the computational model proposed in [78], which describes the emotional learning observed in the mammalian limbic system. The model, which is graphically shown in Figure 2.2, has two main parts: *Amygdala*, which is responsible for immediate learning, and *Orbitofrontal Cortex*, responsible for inhibition of any inappropriate learning happening in the Amygdala.

There are two main inputs to the BELBIC model: *Sensory Inputs* (SI) and *Emotional Signal* (ES). The model output equation (MO) can be described as

$$MO = \sum_i A_i - \sum_i OC_i \tag{2.8}$$

which consists of the difference between Amygdala outputs (A_i) and Orbitofrontal Cortex outputs (OC_i). In this equation, i is the number of sensory inputs.

Amygdala and Orbitofrontal Cortex outputs are calculated by the summation of all their corresponding nodes. Specifically, the output of each node is defined as

$$A_i = V_i SI_i \tag{2.9}$$

$$OC_i = W_i SI_i \tag{2.10}$$

where W_i and V_i are the weights of the Orbitofrontal Cortex and Amygdala, respectively, and SI_i is the i_{th} sensory input. The following equations are utilized for updating V_i and W_i, respectively

$$\Delta V_i = K_v SI_i \max\left(0, ES - \sum_i A_i\right) \tag{2.11}$$

$$\Delta W_i = K_w SI_i \left(MO - ES\right) \tag{2.12}$$

here, K_v and K_w are the learning rates.

Another input considered in the model corresponds to the maximum of all SI. This signal (i.e., A_{th}), which directly comes from the Thalamus to the Amygdala, is described as

$$A_{th} = V_{th} \max\left(SI_i\right) \tag{2.13}$$

Here V_{th} is the weight and the update law is similar to (2.11).

Different approaches have been adopted for tuning the BELBIC parameters. For example, the authors in [79] implemented a particle swarm optimization-based approach. In [80] a Lyapunov-based tuning algorithm was adopted. The authors

in [81] proposed a fuzzy tuning of BELBIC parameters, and successfully utilized this method for controlling a robotic arm. Additionally, the authors in [47] adopted the clonal selection algorithm for obtaining BELBIC parameters, and applied it to the control of a single-link flexible joint manipulator. Trial and error tuning has also shown to be appropriate, since [43], [41], and [76] relied on this method. In this paper, a heuristic approach is adopted for defining the appropriate BELBIC parameters that can significantly reduce the computational complexity.

2.2.3 Control Objectives

Based on the quad rotorcraft model introduced in Subsection 2.2.1, and by employing the BELBIC methodology introduced in Subsection 2.2.2, the objective of this research is to design a set of control signals $\{U_1,U_2,U_3,U_4\}$ for the quad rotorcraft UAS, in such a way that the system is stabilized during flight in all six DoF, even when the system dynamics are uncertain or unknown, and even in the presence of environmental noise/disturbances. Furthermore, the proposed controller should be practically feasible, i.e., should be capable of handling model uncertainty, environmental noises/disturbances and also low computational complexity which makes it appropriate for being implemented in the real-time stabilization of an experimental platform, in our case, the BEBOP Drone from Parrot.

2.3 BELBIC-based Intelligent Control for UAS

2.3.1 System Design

Generally, there are two approaches, so called direct and indirect, for utilizing the cognitive and/or intelligent control. The intelligent and/or cognitive model is utilized as a controller block in the direct methodology, while the intelligent and/or cognitive model is employed for obtaining the controller's parameters in indirect methodology. Here, in our work, we adopted the first method i.e., direct, where a neurobiologically-inspired intelligent methodology based on the computational model of emotional learning in the mammalian brain is used as the controller block.

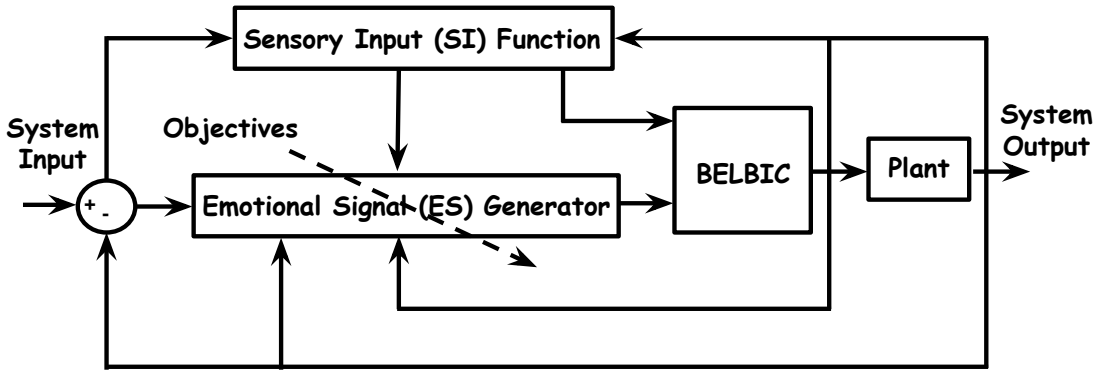

Figure 2.3: BELBIC in the control loop.

Since this research emphasizes the real-time implementation requirement, BELBIC emerges as a natural solution, since its implementation can be accomplished without increasing the complexity of the overall system. Figure 2.3 shows the BELBIC architecture implemented in this work. The figure shows a closed loop configuration consisting of BELBIC block, Emotional signal (ES) generator block, Sensory

inputs (SI) function block and finally a block for the plant. The overall emotional based control concept has been implicitly demonstrated in this architecture, which consists of the critic, learning algorithm, and the action selection mechanism [41].

2.3.2 Emotional Signal and Sensory Input Development

Essentially, BELBIC is an action selection methodology which produces its action based on Emotional signal (ES) and Sensory input (SI). In general, SI and ES can be obtained by using the following expressions

$$SI = G\left(y, e, u, r\right) \tag{2.14}$$

$$ES = F\left(y, e, u, r\right) \tag{2.15}$$

where y is the system output, e is the system error, u is the control effort, and r is the system input. The designer can implicitly decide the control objectives e.g. tracking, flocking, optimal control etc., by choosing the adequate ES. For example, it is possible to choose the ES to reduce the energy expense, to reduce overshoot, or to achieve a better target tracking performance, among others.

Aiming at designing a model-free controller well suited for implementation in real-time systems, the BELBIC design proposed here will focus on improving: (i) target tracking, (ii) disturbance rejection, and (iii) model uncertainty handling. To fulfill these objectives, the ES is designed in such a way that the increase in tracking error will represent a *negative emotion*, e.g., stress, which is then taken as an evidence that the system performance is not satisfactory. Then, for each of the control inputs $\{U_1$,U_2,U_3,$U_4\}$, the SI_l and ES_l, for $l = \{1, \ldots, 4\}$ will be designed as

$$SI_l = K_{l1}e + K_{l2}\int e.dt + K_{l3}\frac{de}{dt} \tag{2.16}$$

$$ES_l = K_{l4}e + K_{l5}\int e.dt + K_{l6}\frac{de}{dt} \tag{2.17}$$

where K_{l1}, K_{l2}, K_{l3}, K_{l4}, K_{l5}, and K_{l6} are positive gains. By assigning different values to these gains, the ES will change its influence on the system behavior. In this work, different values are used for each one of control inputs of the system, i.e., U_l, $l = 1, ..., 4$.

2.3.3 Learning-based Tracking Control

Since multiple performance considerations have to be taken into account all at the same time in tracking control of UAS, it is a very interesting case for using learning-based multi-objective methodologies like BELBIC. Intelligent tracking control of UAS considering the problem of handling the system uncertainties and disturbances, additionally being suitable for real-time implementation due to its control effort reduction capability and low computational complexity, encourages us to take advantage of the computational model of emotional learning in the mammals' limbic system, i.e., BELBIC.

From equations (2.16)-(2.17), the BELBIC-inspired control strategy for stabilization of a quad rotorcraft UAS is defined as

$$\begin{aligned} U_l &= \sum_l V_l.SI_l - \sum_l W_l.SI_l \\ &= \sum_l V_l.\left(K_{l1}e + K_{l2}\int e.dt + K_{l3}\frac{de}{dt}\right) \\ &\quad - \sum_l W_l.\left(K_{l1}e + K_{l2}\int e.dt + K_{l3}\frac{de}{dt}\right) \end{aligned} \tag{2.18}$$

Here, $l = 1, ..., 4$ makes reference to each control input. Considering the results obtained in Theorem 1 and by substituting the Emotional Signal with equation

(2.17) the BELBIC model output for tracking control of a UAS could be obtained as follows:

$$MO = ES_l = K_{l4}e + K_{l5}\int e.dt + K_{l6}\frac{de}{dt} \tag{2.19}$$

which is clearly satisfies our goal of reducing tracking error.

The overall BELBIC-inspired quad rotorcraft control methodology proposed in this paper is summarized as pseudo-code in **Algorithm** 1.

Algorithm 1 : The BELBIC-inspired methodology for controlling a quad rotorcraft.

Initialization:
Set $V_l = 0$, $W_l = 0$, and $V_{th} = 0$, for $i = l, ..., 4$.
Define $ES_l =$ Objective function, for $l = 1, ..., 4$.
  for each iteration $t = t_s$ **do**
    for each control inputs l **do**
      Compute $SI_l = K_{l1}e + K_{l2}\int e.dt + K_{l3}\frac{de}{dt}$
      Compute $ES_l = K_{l4}e + K_{l5}\int e.dt + K_{l6}\frac{de}{dt}$
      Compute $A_l = V_l SI_l$
      Compute $OC_l = W_l SI_l$
      Compute $A_{th} = V_{th}\max(SI_l)$
      Compute $U_l = \sum_l A_l - \sum_l OC_l$
      Update V_l
      Update W_l
      Update V_{th}
    end for
  end for

2.3.4 Stability Analysis

The convergence of the weights of Amygdala (V_l) and Orbitofrontal Cortex (W_l) are presented in Theorem 1.

Theorem 1. *Given the BELBIC design as* (2.16)–(2.19)*, there exists the positive BELBIC tuning parameter, K_v, K_w satisfying*

I. $\left|\left[1 - K_v \left(SI_l\right)^2\right]\right| < 1$

II. $\left|\left[1 - K_w \left(SI_l\right)^2\right]\right| < 1$

such that the BELBIC estimated weights of Amygdala (V_l) and Orbitofrontal Cortex (W_l) converge to desired targets asymptotically.

Proof. **Non-adapting Phase**

Our goal is to investigate the output of the system in non-adapting phase (i.e., when the system completes its learning process) so the equations (2.11) and (2.12) which are the updating rules of Amygdala and Orbitofrontal Cortex, respectively, should be taken into consideration. In addition we make an assumption that the max function in equation (2.11) could be neglected. By substituting (2.9) and (2.10) in equation (2.8) the output of the model could be defined as follows:

$$\begin{aligned} MO &= \sum_l V_l SI_l - \sum_l W_l SI_l \\ &= \sum_l \left(V_l - W_l\right) SI_l \end{aligned} \tag{2.20}$$

$$\begin{aligned} \Delta V_l &= K_v SI_l \left(ES - \sum_l A_l\right) \\ &= K_v SI_l \left(ES - \sum_l V_l SI_l\right) \end{aligned} \tag{2.21}$$

$$\begin{aligned}\Delta W_l =& K_w SI_l \left(MO - ES\right) \\ =& K_w SI_l \left(\sum_l \left(V_l - W_l\right) SI_l - ES\right)\end{aligned} \tag{2.22}$$

When the learning process is completed (i.e., after system completes its learning process) the variations of the weights of Amygdala (ΔV_l) and Orbitofrontal Cortex (ΔW_l) will be equal to zero (i.e., $\Delta V_l = \Delta W_l = 0$). With the assumption of $SI_l \neq 0$ the following holds:

$$\begin{aligned}K_v SI_l \left(ES - \sum_l V_l SI_l\right) =& 0 \\ \Rightarrow \sum_l V_l SI_l =& ES\end{aligned} \tag{2.23}$$

$$\begin{aligned}K_w SI_l \left(\sum_l \left(V_l - W_l\right) SI_l - ES\right) =& 0 \\ \Rightarrow \sum_l \left(V_l - W_l\right) SI_l =& ES \\ \Rightarrow \sum_l W_l SI_l =& 0\end{aligned} \tag{2.24}$$

By substituting (2.23) and (2.24) in equation (2.20) the model output in non-adapting phase will be as follows:

$$MO = \sum_l \left(V_l - W_l\right) SI_l = \sum_l V_l SI_l = ES \tag{2.25}$$

Main Proof

Considering the results obtained in Subsection 2.3.4 the following should be achieved:

$$MO_l \rightarrow ES_l \tag{2.26}$$

Let's considering the V_l^* is the weight of Amygdala for each control input l when the system has been learned and let's $\hat{ES}_l$ be the Emotional Signal for each control input l during the adaptation phase. The following hold:

$$ES_l = V_l^* SI_l \;\; and \;\; \hat{ES}_l = V_l SI_l \tag{2.27}$$

$$\Delta V_l(k) = K_v SI_l \max\left(0, ES_l - \hat{ES}_l\right) \tag{2.28}$$

We will investigate the results of the following two cases:

I. $ES_l - \hat{ES}_l \geq 0$

II. $ES_l - \hat{ES}_l < 0$

Considering the case I., the proof can be achieved as follows:

$$\begin{aligned}
\Delta V_l(k) =& K_v SI_l \max\left(0, ES_l - \hat{ES}_l\right) \\
=& K_v SI_l \left(ES_l - \hat{ES}_l\right) \\
=& K_v SI_l \left(V_l^* SI_l - V_l SI_l\right) \\
=& K_v SI_l \left(V_l^* - V_l\right) SI_l \\
=& K_v SI_l \tilde{V}_l SI_l \\
=& K_v \tilde{V}_l \left(SI_l\right)^2
\end{aligned} \tag{2.29}$$

where $\tilde{V}_l = V_l^* - V_l$.

$$\begin{aligned}
V_l\left(k+1\right) =& V_l\left(k\right) + \Delta V_l(k) \\
\tilde{V}_l\left(k+1\right) =& V^* - V_l\left(k\right) - \Delta V_l(k) \\
=& \tilde{V}_l\left(k\right) - K_v \tilde{V}_l \left(SI_l\right)^2 \\
=& \left[1 - K_v \left(SI_l\right)^2\right] \tilde{V}_l\left(k\right)
\end{aligned} \tag{2.30}$$

Considering the case II., it is obvious that when $ES_l - \hat{ES}_l < 0$ the max function in equation (2.28) will force the adaptation in Amygdala to stop and the following hold:

$$\begin{aligned} \Delta V_l(k) =& 0 \\ V_l\left(k+1\right) =& V_l\left(k\right) \\ \tilde{V}_l\left(k+1\right) =& \tilde{V}_l\left(k\right) \end{aligned} \tag{2.31}$$

The proof can be achieved as follows:

$$\begin{aligned} \Delta W_l(k) =& K_w SI_l\left(MO_l - ES_l\right) \\ =& K_w SI_l\left(V_l SI_l - W_l SI_l - V_l^* SI_l\right) \\ =& K_w SI_l\left(-\left(V_l^* - V_l\right) SI_l - W_l SI_l\right) \\ =& K_w SI_l\left(\left(-\tilde{V}_l - W_l\right) SI_l\right) \\ =& -K_w \tilde{W}_l \left(SI_l\right)^2 \end{aligned} \tag{2.32}$$

where $\tilde{V}_l = V_l^* - V_l$ and $\tilde{W}_l = \tilde{V}_l + W_l$.

$$\begin{aligned} W_l\left(k+1\right) =& W_l\left(k\right) + \Delta W_l(k) \\ \tilde{W}_l\left(k+1\right) =& \tilde{V}_l\left(k+1\right) + W_l\left(k+1\right) \\ =& \tilde{V}_l\left(k\right) + W_l\left(k\right) + \Delta W_l(k) \\ =& \tilde{W}_l\left(k\right) - K_w \tilde{W}_l \left(SI_l\right)^2 \\ =& \left[1 - K_w \left(SI_l\right)^2\right] \tilde{W}_l\left(k\right) \end{aligned} \tag{2.33}$$

□

Theorem 2. *(Closed-loop Stability): Given the initial UAS state $x(0)$ and the BELBIC estimated weights of Amygdala ($V_l(0)$) and Orbitofrontal Cortex ($W_l(0)$) be bounded in the set $\mathbf{\Lambda}$. Let the BELBIC be tuned and estimated control policy be given as* (2.21), (2.22) *and* (2.18)

respectively. Then, there exists positive constants, K_v, K_w, satisfying Theorem 1 such that UAS state, $x(t)$ and BELBIC weights estimation errors are all asymptotically stable.

Proof. Let's consider the u_s is a stable controller for the following system:

$$\dot{x} = f(x) + g(x)u_s \tag{2.34}$$

There is a Lyapunov function $L_s(x)$ which guarantees the stability of the whole system:

$$L_s(x) = \frac{1}{2}x^T x$$

Taking the first derivative, we have:

$$\begin{aligned} \dot{L}_s(x) &= x^T \dot{x} \\ &= x^T[f(x) + g(x)u_s] \\ &\leq -lx^T x \quad , \ l > 0 \end{aligned} \tag{2.35}$$

To provide the stability analysis of the actual system, let's consider the u_a is an actual controller for the following system:

$$\dot{x} = f(x) + g(x)u_a \tag{2.36}$$

where u_a is as follows:

$$u_a = u_s + \tilde{u} \tag{2.37}$$

and $\tilde{u}$ is the controller which is given by the BELBIC model output MO. Considering the Lyapunov function $L_{MO}(x)$, the following is obtained:

$$L_{MO}(x) = A(\widetilde{MO})^2$$

Taking the first derivative, we have:

$$\begin{aligned}\dot{L}_{MO}(x) &= A(\widetilde{MO})(\dot{\widetilde{MO}}) \\ &\leq -A(\widetilde{MO})^2 \\ &\leq -A(\tilde{u})^2 \end{aligned} \tag{2.38}$$

Considering the Lyapunov function $L_a(x)$, the stability proof of overall system is as follows:

$$L_a(x) = \frac{1}{2}x^T x$$

Taking the first derivative, we have:

$$\begin{aligned}\dot{L}_a(x) &= x^T \dot{x} \\ &= x^T[f(x) + g(x)u_a] \\ &= x^T[f(x) + g(x)(u_s + \tilde{u})] \\ &= x^T[f(x) + g(x)u_s + g(x)\tilde{u}] \\ &= x^T[f(x) + g(x)u_s] + x^T g(x)\tilde{u} \\ &\leq -lx^T x + x^T g(x)\tilde{u} \\ &\leq -lx^T x + \frac{l}{2}x^T x + \frac{2}{l}(g(x)\tilde{u})^2 \\ &\leq -\frac{l}{2}x^T x + \frac{2}{l}(g(x)\tilde{u})^2 \\ &\leq -\frac{l}{2}x^T x - A(\tilde{u})^2 \end{aligned} \tag{2.39}$$

□

Remark 1. *Based on the BELBIC theory [41] and* (2.18)*, the intelligent tracking control of UAS can be obtained while the estimated weights of Amygdala (V_l) and Orbitofrontal Cortex (W_l) are converging to desired targets. According to Theorem 1, estimated weights converge to desired targets asymptotically. Therefore, the designed BELBIC input U_l* (2.18) *converges to intelligent tracking control of UAS asymptotically.*

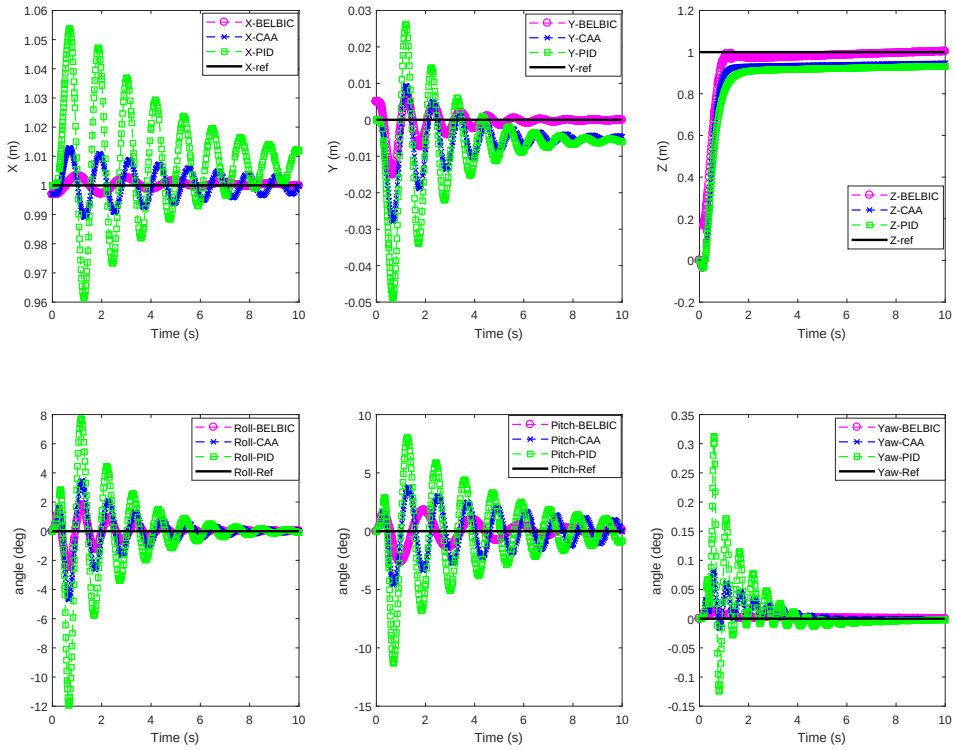

Figure 2.4: Altitude Control Results: In this experiment, the UAS should vertically take off and reach an altitude of 1 meter, while keeping its X, Y, θ, ϕ, and ψ states equal to zero. Outputs of the controlled system: the BELBIC-inspired control is shown in Magenta color, while the CAA controller in Blue color and the PID controller in Green color.

2.4 Simulation Results

A set of numerical simulations are provided in order to demonstrate the effectiveness and performance of the BELBIC-inspired methodology. Additionally, a comparison between the proposed approach and two different control strategies is provided, where it is possible to observe the significant improvement in control performance obtained by utilizing BELBIC.

These simulation results consider four different scenarios: stable vertical take off is presented in Subsection 2.4.1, while trajectory tracking is addressed in Subsection 2.4.2. Additionally, disturbance rejection results are shown in Subsection 2.4.3,

and system parameters uncertainty handling is investigated in Subsection 2.4.4. In all the scenarios, a quad rotorcraft model is adopted, with dynamics as presented in Subsection 2.2.1. Initial velocities are selected equal to zero, and initial position equal to $(X_0 = 1, Y_0 = 0, Z_0 = 0)$. The quad rotorcraft parameters introduced in [43] are used through all the simulations. The remaining parameters are introduced in each subsection, as needed.

2.4.1 Stable Vertical Take-off

In this study, a stable vertical take-off task is addressed. Starting from its initial position, the UAS should vertically take off and reach an altitude of 1 meter, while keeping its X, Y, θ, ϕ, and ψ states equal to zero. For comparison purposes, two similar experiments were performed, but using the PID controller introduced in [77] and Control Allocation Approach (CAA) introduced in [82] instead of the BELBIC controller. Figure 2.4 shows the output of the controlled system by means of the PID controller, Control Allocation Approach (CAA), and the BELBIC-inspired controller. From this Figure, it is possible to observe that the BELBIC-inspired controller outperforms the other controllers' performance.

2.4.2 Trajectory Tracking

In this study, trajectory tracking of a quad rotorcraft is addressed. Here, a 5m$\times$5m square shaped reference is assigned for the system to follow, as shown in Figure 2.5. The scenario is as follows. First, the quad rotorcraft starts by taking off, and reaches an altitude of 1m. After this, it follows the squared trajectory with a

speed of 1 m/s. Then, 1 meter before reaching its starting position, the platform increases its height for another 1 meter, and then flies forward for another 2 meter.

For comparison purposes, in addition to the BELBIC-inspired strategy, two similar experiments were performed but using the PID controller introduced in [77] and Control Allocation Approach (CAA) introduced in [82]. Figures 2.6 and 2.7 show the aircraft (X, Y, Z)-positions and (Roll, Pitch, Yaw)-angles, as controlled by the PID, Control Allocation Approach (CAA), and the BELBIC-inspired controller, respectively. Similar to the previous scenario, it is possible to observe that the BELBIC-inspired controller outperforms the other controllers performance.

2.4.3 Disturbance Rejection

In this study, the performance of the proposed strategy when facing disturbances is investigated. The disturbance comes in the form of a torque of 2 N.m applied to the quad rotorcraft pitch angle. For comparison purposes, two similar experiments were performed, but using the PID controller introduced in [77] and Control Allocation Approach (CAA) introduced in [82] instead of the BELBIC controller. Figure 2.8 shows the pitch angle and rear motor control effort of the quad rotorcraft generated when using the PID controller, CAA controller, and the BELBIC-inspired controller. From this Figure, it is observed that the BELBIC-inspired controller is able to reject the disturbance effectively, and that the obtained behavior is at least 10 times better than the behavior obtained from the PID controller.

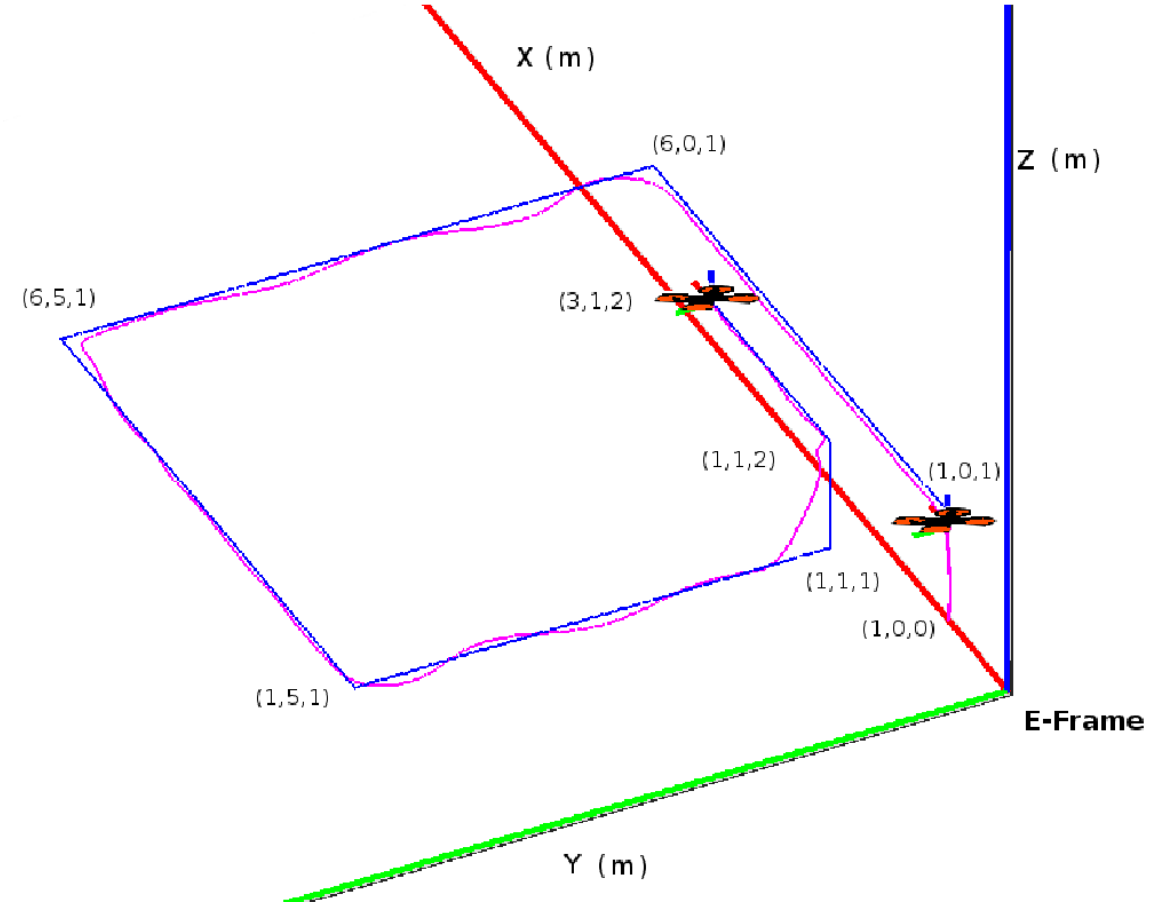

Figure 2.5: Trajectory Tracking of a Quad Rotorcraft BELBIC-inspired control (Magenta), Reference (Blue). A 5m$\times$5m square shaped reference is assigned for the UAS to follow. The figure shows the UAS hovering at $(1, 0, 1)$ above the starting position, and also at the end position $(3, 1, 2)$ of the assigned trajectory.

2.4.4 System Parameters Uncertainty Handling

In this study, handling uncertainty in the system parameters is considered. The same reference trajectory as in Subsection 2.4.2 is assigned to the system. All the parameters of the quad rotorcraft are changed by ± 10 percent of their original values. For comparison purposes, in addition to the BELBIC-inspired strategy, two similar experiments were performed but using the PID controller introduced in [77] and Control Allocation Approach (CAA) introduced in [82]. The objective

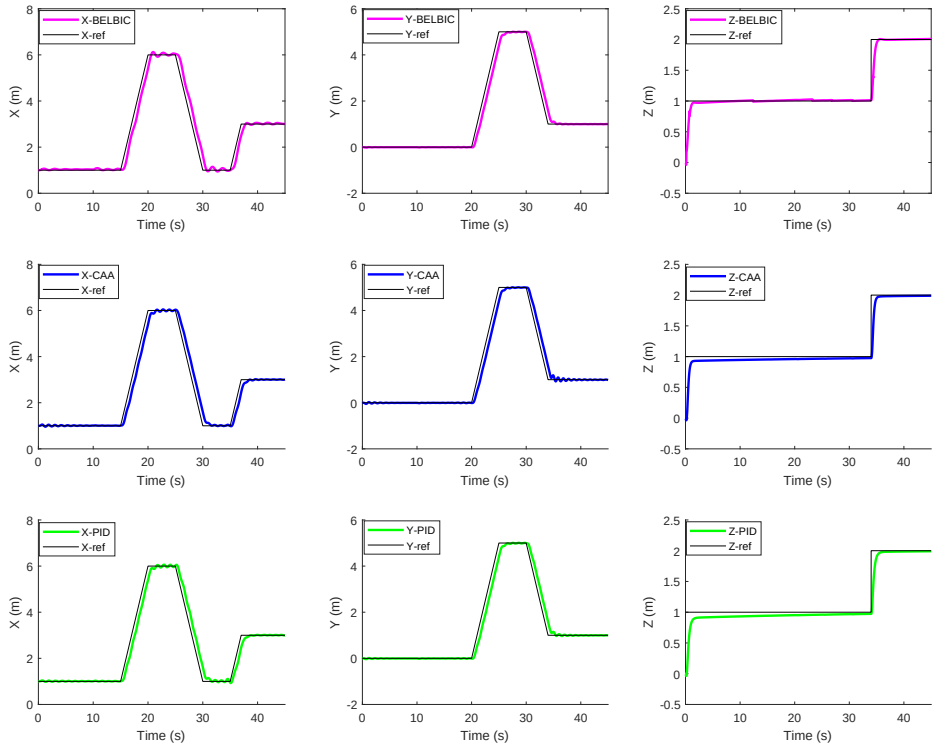

Figure 2.6: The output of the system for intelligent tracking control in X-axis, Y-axis, and Z-axis, from left to right, respectively. The BELBIC-inspired control is in the top row and in Magenta color. The outputs of the CAA controller are plotted on the middle row and in Blue color. The outputs of the PID controller are plotted on the bottom row and in Green color. In all cases, the reference signal is shown in Black color.

here is to verify if all strategies are able to track the desired trajectory, taking into account that all controller settings remain the same as the settings used in Subsection 2.4.2. In other words, no additional tuning of the controllers was performed for adapting to the new system parameters. Figures 2.9 and 2.10 show the quad rotorcraft position and attitude angles, as controlled by the PID controller, CAA controller, and the BELBIC-inspired controller, respectively. It is possible to observe that both CAA and PID controllers are not capable of stabilizing the system with uncertainties, while the BELBIC-inspired strategy is able to successfully control the system.

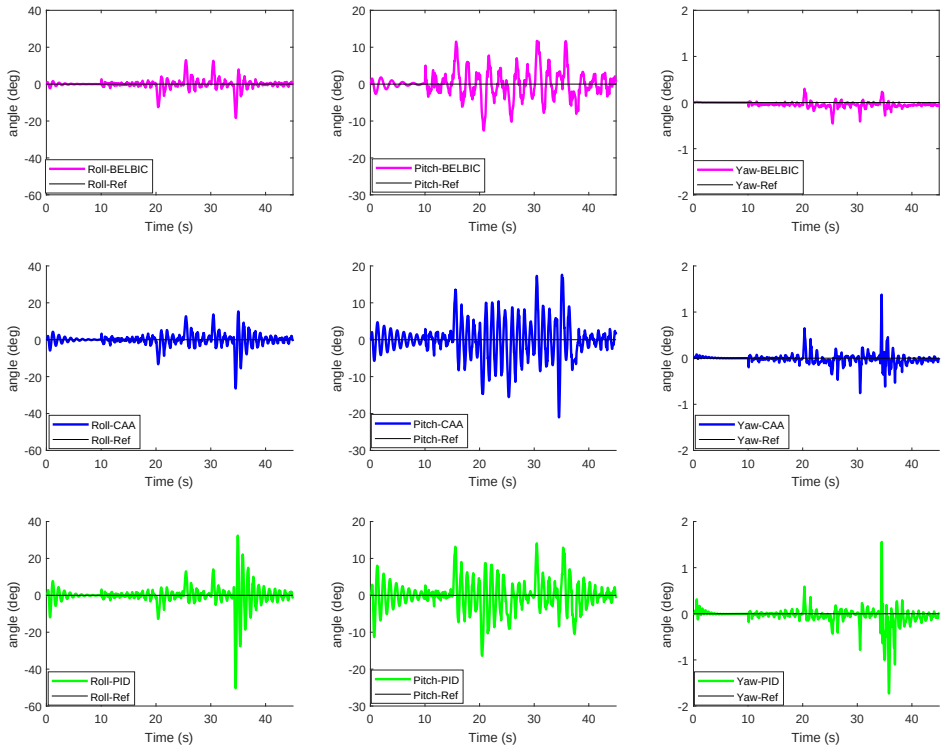

Figure 2.7: The output of the system for intelligent tracking control for Roll, Pitch, and Yaw angles, from left to right respectively. The BELBIC-inspired control is plotted in the top row in Magenta color. The outputs of the CAA controller are plotted in the middle row, in Blue color. The outputs of the PID controller are plotted in the bottom row, in Green color. For all cases, the reference signal is plotted in Black.

2.5 Experimental Results

This section presents experimental results showing the performance of the BELBIC-inspired controller for stabilizing a quad rotor UAS during autonomous flights.

2.5.1 Experimental Testbed

The platform implemented for validation of the proposed algorithm is available at the Unmanned Systems Laboratory (USL) from the University of Nevada - Reno [83].

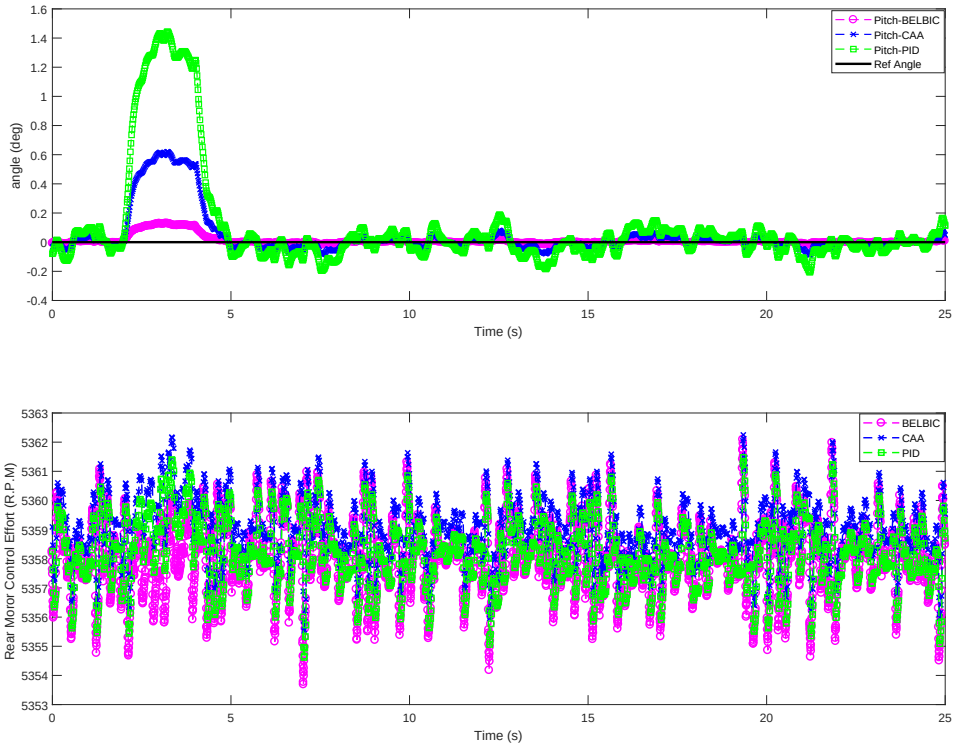

Figure 2.8: Control signal and pitch angle performance: the pitch angle (top plot) and rear motor control effort (bottom plot) of the quad rotorcraft when facing a torque of 2 N.m as disturbance to its pitch angle. The BELBIC-inspired control is shown in Magenta. The CAA controller is shown in Blue. The PID controller is shown in Green, and the reference signal in Black.

The Base Station of this testbed runs Ubuntu 14.04 OS, the Robot Operating System (ROS) environment, and Matlab. The proposed algorithms were coded in C/C++, and were executed at 100Hz. The aerial robot corresponds to a Bebop drone manufactured by Parrot.

The 3-dimensional position of the UAS is obtained by means of a Motion Capture System (MCS) manufactured by OptiTrack. The information provided by the MCS is reported to the OptiTrack Interface PC by means of a Gigabyte Ethernet connection. Next, this information is sent to the Base Station PC by means of an

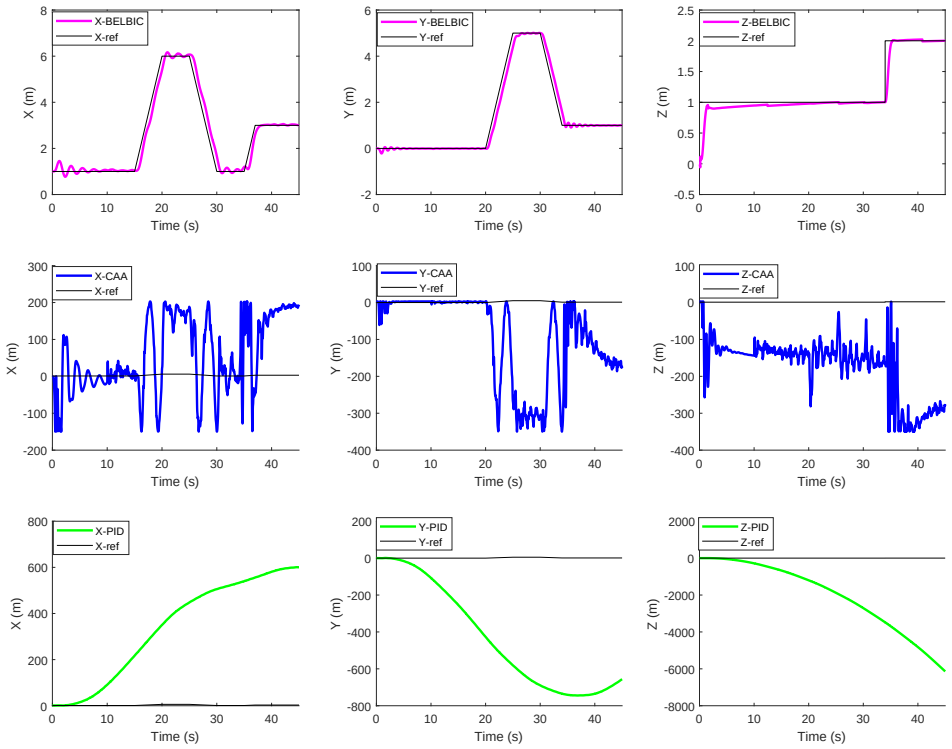

Figure 2.9: The figures show the output of the system for intelligent tracking control in presence of model uncertainty in X-axis, Y-axis, and Z-axis, from left to right, respectively. Outputs of the controlled system: the BELBIC-inspired control is in the top row in Magenta color. The CAA controller is on the middle row in Blue color. The PID controller is on the bottom row in Green color. In all cases the reference is plotted in Black color.

Ethernet connection. The Base Station computer uses this information to execute the BELBIC algorithm and to calculate the control signals, which are then sent to the Bebop platform by means of a WiFi link. The ultimate goal of this experimental application is to maintain a satisfactory flight of the drone, even when the model of the UAS is not known, and when external factors affect the performance of the aerial vehicle. The left hand side (LHS) of Figure 2.11 shows the data flow proposed for the implementation of the BELBIC algorithm. The right hand side (RHS) of Figure 2.11 shows the experimental evaluation of the proposed algorithm on a

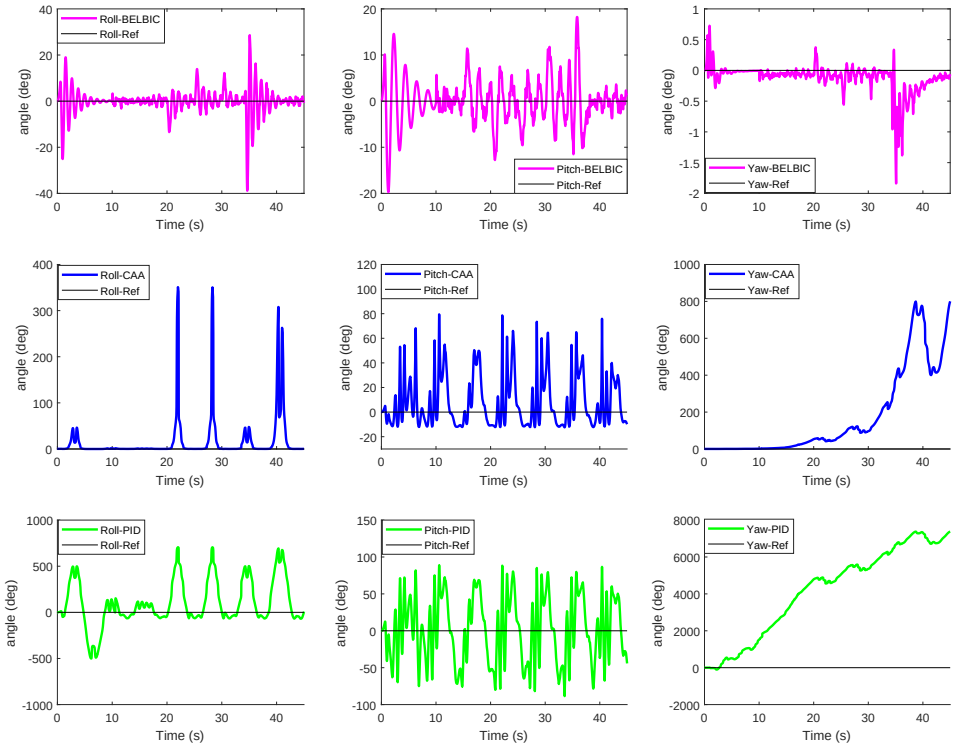

Figure 2.10: The output of the system for intelligent tracking control in presence of model uncertainty for Roll, Pitch, and Yaw angles, from left to right, respectively. The outputs of the controlled system: the BELBIC-inspired control is in the top row in Magenta color. The CAA controller is in the middle row in Blue color. The PID controller is in the bottom row in Green color. The reference signal is plotted in Black color.

Bebop drone.

2.5.2 Real-time Experiments

A set of experimental results demonstrate the satisfactory performance of the proposed method. In this experiment, the attitude control of a quad rotorcraft is presented. Starting from an initial position in the ground, the UAS executes a take off,

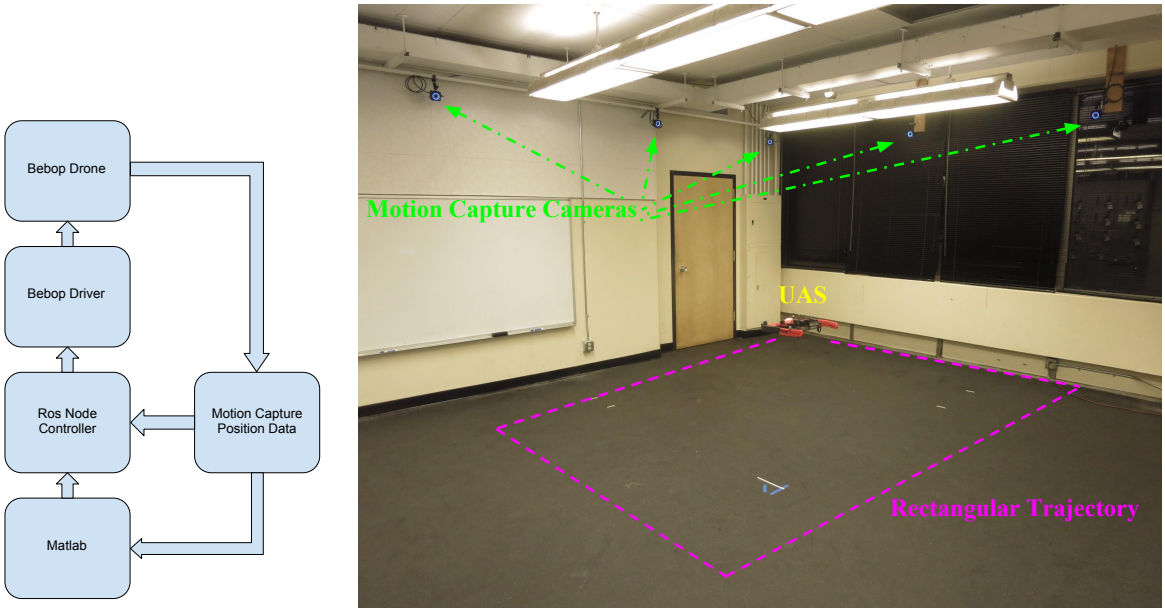

Figure 2.11: Data flow showing the implementation of the BELBIC-inspired algorithm (left); the experimental evaluation of the proposed algorithm on a Bebop drone manufactured by Parrot (right).

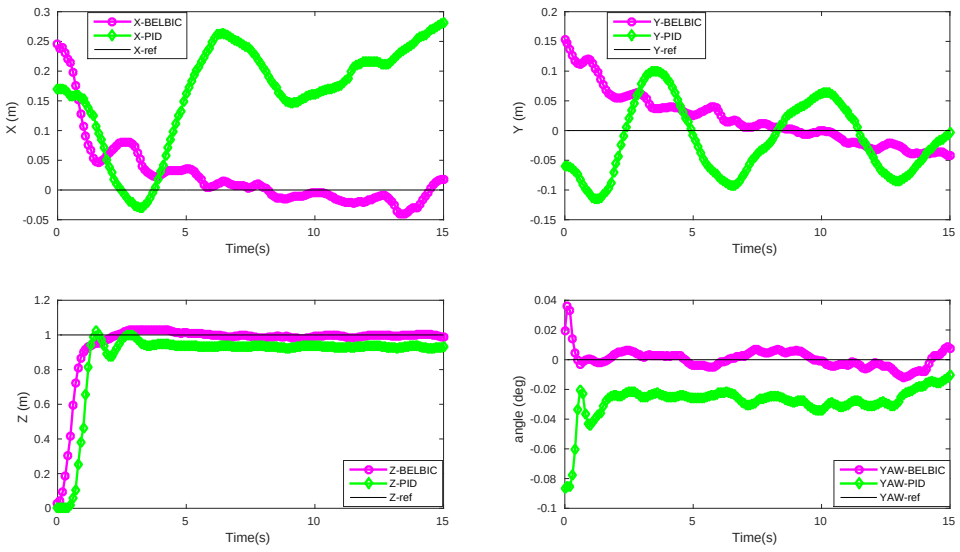

Figure 2.12: Outputs of the controlled system, BELBIC-inspired control (Magenta), conventional PID controller (Green) Reference (Black). Starting from an initial position in the ground, the UAS executes a take off, and reaches an altitude of 1 meter, while keeping its X, Y, θ, ϕ, and ψ states all equal to zero.

and reaches an altitude of 1 meter, while keeping its X, Y, θ, ϕ, and ψ states all equal to zero. For comparison purposes, a similar experiment was performed, but using the conventional PID controller instead of the BELBIC-inspired controller. Figure 2.12 shows the performance of the controlled system, when using the conventional PID controller, and when using the proposed BELBIC-inspired controller. Similar to the results obtained in numerical simulations, it is possible to observe that the BELBIC-inspired controller outperforms the conventional PID controller performance.

Similar to this experiment, it is also possible to investigate the tracking problem addressed in Subsection 2.4.2 which is not included in this paper due to brevity.

2.6 Conclusions

This paper addressed the problem of stabilizing the full 6 DoF of a quad rotorcraft UAS and also intelligent tracking control of a UAS subjected to unknown system dynamics and external disturbances. The low-computational model-free BELBIC, a neurobiologically-motivated intelligent controller, was adopted in order to design and experimentally validate a novel UAS control methodology. The numerical and experimental flight results, which considered uncertainty and disturbances, demonstrated the effectiveness, applicability, and superior performance of the BELBIC-inspired controller, when compared to conventional model-based control methods. In addition, the convergence analysis of the proposed approach has been studied.

Future work will consider the implementation of a BELBIC-inspired intelligent

control strategy for addressing the task of UAS-based autonomous transportation of loads with uncertain characteristics.

Chapter 3

A Biologically-Inspired Reinforcement Learning based Intelligent Distributed Flocking Control for Multi-Agent Systems in Presence of Uncertain System and Dynamic Environment

In this paper, we investigate the real-time flocking control of Multi-Agent Systems (MAS) in presence of system uncertainties and dynamic environment. To handle the impacts from system uncertainties and dynamic environment, a novel

reinforcement learning technique, which is appropriate for real-time implementation, has been integrated with multi-agent flocking control in this paper. The Brain Emotional Learning Based Intelligent Controller (BELBIC) is a biologically-inspired reinforcement learning based controller relying on a computational model of emotional learning in the mammalian limbic system. The learning capabilities, multi-objective properties, and low computational complexity of BELBIC make it a very promising learning technique for implementation in real-time applications. Firstly, a novel brain emotional learning based flocking control structure is proposed. Then, the real-time update laws are developed to tune the emotional signals based on the real-time operation data. It is important to note that this data-driven reinforcement learning approach relaxes the requirement for system dynamics and effectively handle the uncertain impacts of the environment. Using the tuned emotional signals, the optimal flocking control can be obtained. The Lyapunov analysis has been used to prove the convergence of the proposed design. The effectiveness of the proposed design is also demonstrated through numerical and experimental results based on the coordination of multiple Unmanned Aircraft Systems (UAS) platforms.

3.1 Introduction

During the past decade, diverse research communities have developed several advanced control strategies for coordination of Multi-Agent Systems (MAS), see for example [1, 3–12] and the references therein. In most of these MAS control methods, *flocking* problem have been formulated and investigated thoroughly. Flocking is the collective motion of a large number of self-propelled entities exhibited by many living beings such as birds, fish, bacteria, and insects [15]. Flocking is also

considered as an emergent behavior, which is caused by a number of simple rules followed by each agent, and that does not require any central coordination.

The seminal work in [7] introduced three basic rules for simulating the flocking behavior, specifically, *separation*, *alignment*, and *cohesion*. In recent years, there has been a surge of interest among researchers to improve the flocking behavior of MAS. One can mention, for example, adaptive flocking control approaches for dealing with varying and noisy environments [16], [17], robust flocking controllers to handle model uncertainty [18, 19], and flocking control methods with the capability of disturbance handling [20].

Related Work

Several critical aspects should be considered in the real-time implementation of control strategies on mobile robotic platforms, for example, model uncertainty, disturbances, energy expenses and actuator saturation. Diverse research efforts have been proposed aiming at addressing the issues arising from these practical and harsh conditions. Considering model uncertainty and disturbances, distributed tracking and estimation of MAS with uncertain dynamics has been presented in [27]. In [28], the problem of robust consensus tracking for MAS with disturbances and unmodeled dynamics has been studied. In addition, neural adaptive flocking of MAS has been addressed in [29–31]. More recently, authors in [32] investigated the application of Q-learning in leader-follower based flocking with small fixed-wing UAVs.

Also, to tackle the energy expenses and actuator saturation related problems, the authors in [33] presented a decentralized approach to perform formation ma-

neuvers by groups of mobile robots taking into account the actuator saturation. Closely related, a flocking control with constraints is proposed in [34]. Swarm aggregation of MAS with actuator saturation has been addressed in [35,36]. Furthermore, a leader-following tracking problem for MAS with a varying-velocity leader and input saturation was investigated in [37]. In [38], the authors introduced a decentralized connectivity maintenance strategy for mobile networks with constrained inputs. Recently, an energy function-based approach for estimating the required control force for network connectivity preservation and collision avoidance was presented in [39].

In general, flocking strategies available in the literature are addressing, among others i.e., (i) the optimization of the MAS flocking control, (ii) the robustness for dealing with the dynamic environment, and (iii) the capability of dealing with MAS uncertainty. However, most of these recent approaches are not designed for having multi-objective properties i.e., they only considered solving one of the aforementioned problems. Indeed, when multiple goals are targeted in parallel to the flocking problem, the computational complexity of the overall problem is not suitable for real-time implementation. Additionally, most of the existing works need the knowledge of system dynamics and not applicable in presence of dynamic environment.

Therefore, to overcome this deficiency, biologically-inspired methods have been increasingly employed to solve complex computational problems. Brain Emotional Learning Based Intelligent Controller (BELBIC) is one of the most promising approaches that adopts the learning model developed in [40] in order to mimic the functionalities of the brain that are known to produce emotion, i.e., the *amygdala*, *orbitofrontal cortex*, *thalamus*, and *sensory input cortex*. Strategies based on BELBIC

have shown to be a very promising solution in terms of robustness and uncertainty handling [41–43] due to the flexibility of its architecture. BELBIC has two main inputs: *Sensory Inputs* (SI) and *Emotional Signal* (ES). Properly adjusting both SI and ES, makes this controller an appealing strategy for addressing real-time applications [41]. Furthermore, BELBIC architecture has the computational complexity on the order of $O(n)$ [44] that is much smaller and better for real-time implementation compared with other existing learning-based intelligent control methods.

Main Contributions

The main contribution of this paper is to utilize the computational model of emotional learning in the mammalian limbic system, i.e., BELBIC, for developing and experimental validation of a novel intelligent flocking control method for practical MAS in presence of uncertainty system dynamics and dynamic environment. To the best of the authors' knowledge, this is the first time that BELBIC is implemented for accomplishing MAS flocking in an intelligent and practical manner. The learning capability of the proposed approach enhances the flocking strategy, which is very useful when dealing with uncertainty due to the fact that it does not depend on the system model. Furthermore, the computational cost of implementing this method in a real-time application is feasible. The proposed BELBIC design provides a controller with multi-objective properties i.e., control optimization, uncertainty handling, and noise/disturbance rejection while keeping the system complexity in a practically achievable limit. The main objective is then the design of a feasible intelligent controller that is able of keeping the flocking performance as satisfactory as possible, even in presence of system uncertainties and dynamic environment, in terms of formation control, obstacles avoidance, and target track-

ing, and also that can be practically applied to real-time systems. Thus, energy expense and the generation of control signals appropriate for real-time implementation are addressed. Moreover, we proved that our proposed methodology guarantees the designed control signals converge to ideal optimal flocking strategies. A set of numerical simulations and hardware experimental tests are provided in order to demonstrate the effectiveness of the proposed approach. Additionally, a comparison between the proposed approach and two different flocking strategies is provided, where it is possible to observe the improved flocking strategy attained with BELBIC.

The rest of the paper is organized as follows. Section 3.2 presents the problem formulation and some preliminaries about flocking control model and BELBIC. Our main contribution is introduced in Section 3.3, which consists of a flocking control strategy based on BELBIC (Subsection 3.3.1) and its learning weights convergence analysis (Subsection 3.3.3). Section 3.4 and Section 3.5 present numerical simulation and experimental results, respectively. The conclusion of the paper and future directions of our work are provided in Section 3.6.

3.2 Problem Formulation and Preliminaries

3.2.1 Flock Modelling

Considering n agents with second order dynamics moving in an m dimensional space $(m = 2, 3)$, the motion of MAS can be represented as

$$\begin{cases} \dot{q}_i = p_i \\ \dot{p}_i = u_i \quad , i = 1, 2, ..., n \end{cases} \tag{3.1}$$

where $\{u_i, q_i, p_i\} \in \mathbb{R}^{\mathrm{m}}$ are control input, position, and velocity of the agent i, respectively. A dynamic graph $\mathcal{G}(\upsilon, \varepsilon)$ consisting of a set of vertices υ and edges ε is represented by $\upsilon = \{1, 2, ..., n\}$, $\varepsilon \subseteq \{(i, j) : i, j \in \upsilon, j \neq i\}$. Each agent is represented by a vertex, and an individual edge represents a communication link between a pair of agents. Consider the neighborhood set of agent i defined as

$$N_i^{\alpha} = \{j \in \upsilon_{\alpha} : \| q_j - q_i \| < r, j \neq i\} \tag{3.2}$$

with r, a positive constant, being the range of interaction between agents i and j, and the Euclidean norm in $\mathbb{R}^{\mathrm{m}}$ is expressed as $\| \cdot \|$. To describe the geometric model of the flock, i.e., the α-lattice [1], the following set of algebraic conditions should be solved

$$\| q_j - q_i \| = d \quad \forall j \in N_i^{\alpha} \tag{3.3}$$

where d is a positive constant representing the distance between two neighbors i and j. At $q_i = q_j$ the equation (3.3) causes a singularity for the collective potential function [1]. Equation (3.3) can be rewritten to resolve the aforementioned problem as

$$\| q_j - q_i \|_{\sigma} = d_{\alpha} \quad \forall j \in N_i^{\alpha} \tag{3.4}$$

where $d_\alpha = \| d\|_\sigma$, and $\| \cdot\|_\sigma$ is the σ-norm expressed by $\| z \|_\sigma = \frac{1}{\epsilon}[\sqrt{1+\epsilon\| z\|^2} - 1], \epsilon > 0$. Here ϵ is a positive constant. The σ-norm is a map from $\mathbb{R}^m$ to $\mathbb{R} \geq 0$ for a vector z. While the Euclidean norm $\| z\|$ is differentiable everywhere except at $z = 0$, the new map $\| z\|_\sigma$ is differentiable everywhere.

From the above constraints, a smooth collective potential function can be given as

$$V(q) = \frac{1}{2}\sum_i \sum_{j \neq i} \psi_\alpha(\| q_j - q_i\|_\sigma)$$

where $\psi_\alpha(z)$ is a smooth pairwise potential function defined as $\psi_\alpha(z) = \int_{d_\alpha}^{z} \phi_\alpha(s)ds$, with $\phi_\alpha(z) = \rho_h(z/r_\alpha)\phi(z - d_\alpha)$, $\phi(z) = \frac{1}{2}[(a+b)\sigma_1(z+c) + (a-b)]$, and $\sigma_1(z) = z/\sqrt{1+z^2}$. Additionally, $\phi(z)$ is an uneven sigmoidal function with parameters $0 < a \leq b$, $c = |a-b|/\sqrt{4ab}$ to guarantee that $\phi(0) = 0$. The term $\rho_h(z)$ is a scalar bump function that smoothly varies between [0,1]. Inspired from [1], a possible choice for defining $\rho_h(z)$ is as:

$$\begin{cases} 1, & z \in [0, h) \\ \frac{1}{2}\left[1 + \cos\left(\pi \frac{(z-h)}{(1-h)}\right)\right], & z \in [h, 1] \\ 0, & \text{otherwise} \end{cases} \tag{3.5}$$

Recall to [1], the flocking control can be developed as

$$u_i = u_i^\alpha + u_i^\beta + u_i^\gamma \tag{3.6}$$

This design ensures that MAS avoids obstacles while making all agents to form a lattice configuration, i.e., the α-lattice [1]. The algorithm consists of three main terms: u_i^α which is the interaction component between two α-agents, u_i^β that is the interaction component between the α-agent and an obstacle (named the β-agent),

and u_i^γ being a goal component which consists of a distributed navigational feedback term. Specifically, these terms, u_i^α, u_i^β, and u_i^γ, are defined as

$$u_i^\alpha = c_1^\alpha \sum_{j \in N_i^\alpha} \phi_\alpha(\| q_j - q_i \|_\sigma) \mathbf{n}_{i,j} + c_2^\alpha \sum_{j \in N_i^\alpha} a_{ij}(q)(p_j - p_i) \tag{3.7}$$

$$u_i^\beta = c_1^\beta \sum_{k \in N_i^\beta} \phi_\beta(\| \hat{q}_{i,k} - q_i \|_\sigma) \hat{\mathbf{n}}_{i,k} + c_2^\beta \sum_{k \in N_i^\beta} b_{i,k}(q)(\hat{p}_{i,k} - p_i) \tag{3.8}$$

$$u_i^\gamma = -c_1^\gamma \sigma_1(q_i - q_r) - c_2^\gamma (p_i - p_r) \tag{3.9}$$

where c_1^α, c_2^α, c_1^β, c_2^β, c_1^γ, and c_2^γ are all positive constants which are the corresponding weights of the three main flocking control functions, i.e., collision avoidance, obstacle avoidance and goal [1]. Next, the pair (q_r, p_r) is defined as virtual leader of the flock, i.e., the γ-agent which can be represented as

$$\begin{cases} \dot{q}_r = p_r \\ \dot{p}_r = f_r(q_r, p_r) \end{cases} \tag{3.10}$$

The terms $\mathbf{n}_{i,j}$ and $\hat{\mathbf{n}}_{i,k}$ are vectors given as

$$\mathbf{n}_{i,j} = \frac{q_j - q_i}{\sqrt{1 + \epsilon \| q_j - q_i \|^2}} \qquad \hat{\mathbf{n}}_{i,k} = \frac{\hat{q}_{i,k} - q_i}{\sqrt{1 + \epsilon \| \hat{q}_{i,k} - q_i \|^2}}$$

Then, the terms $b_{i,k}(q)$ and $a_{ij}(q)$ are the elements of the heterogeneous adjacency matrix $B(q)$ and spatial adjacency matrix $A(q)$, which are described respectively as follows: $b_{i,k}(q) = \rho_h(\| \hat{q}_{i,k} - q_i \|_\sigma / d_\beta)$ and $a_{ij}(q) = \rho_h(\| q_j - q_i \|_\sigma / r_\alpha) \in [0, 1], j \neq i$. It is important to note that, $r_\alpha = \| r \|_\sigma$, $a_{ii}(q) = 0$ for all i and q, $d_\beta = \| d' \|_\sigma$, and $r_\beta = \| r' \|_\sigma$. Here d' is a positive constant representing the distance between

an α-agent with obstacles. The term $\phi_\beta(z)$ is a repulsive action function which is defined as $\phi_\beta(z) = \rho_h(z/d_\beta)(\sigma_1(z - d_\beta) - 1)$. Now, it is possible to define the set of β-neighbors of an α-agent i in a similar way to equation (3.2) as

$$N_i^\beta = \{k \in \upsilon_\beta : \| \hat{q}_{i,k} - q_i \| < r^{'}\}$$

with the positive constant r' being the range of interaction of an α-agent with obstacles.

In summary, the multi-agent system considered in this research consists of a group of holonomic quad rotorcraft UAS [83]. The objectives of the agents are to avoid obstacles, avoid collisions, and tracking the target while making all agents to form a lattice configuration, i.e., the α-lattice [1]. Each agent will exchange its operation information (e.g. position, velocity, etc.) with their neighbors. This connectivity has been demonstrated through a dynamic graph. A dynamic graph consisting of a set of vertices and edges is considered to represent the system. Each agent is represented by a vertex, and an individual edge represents a communication link between a pair of agents.

The flocking algorithm consists of three main terms: (i) the interaction component between two α-agents, (ii) the interaction component between the α-agent and an obstacle (named the β-agent), and (iii) a goal component which consists of a distributed navigational feedback term.

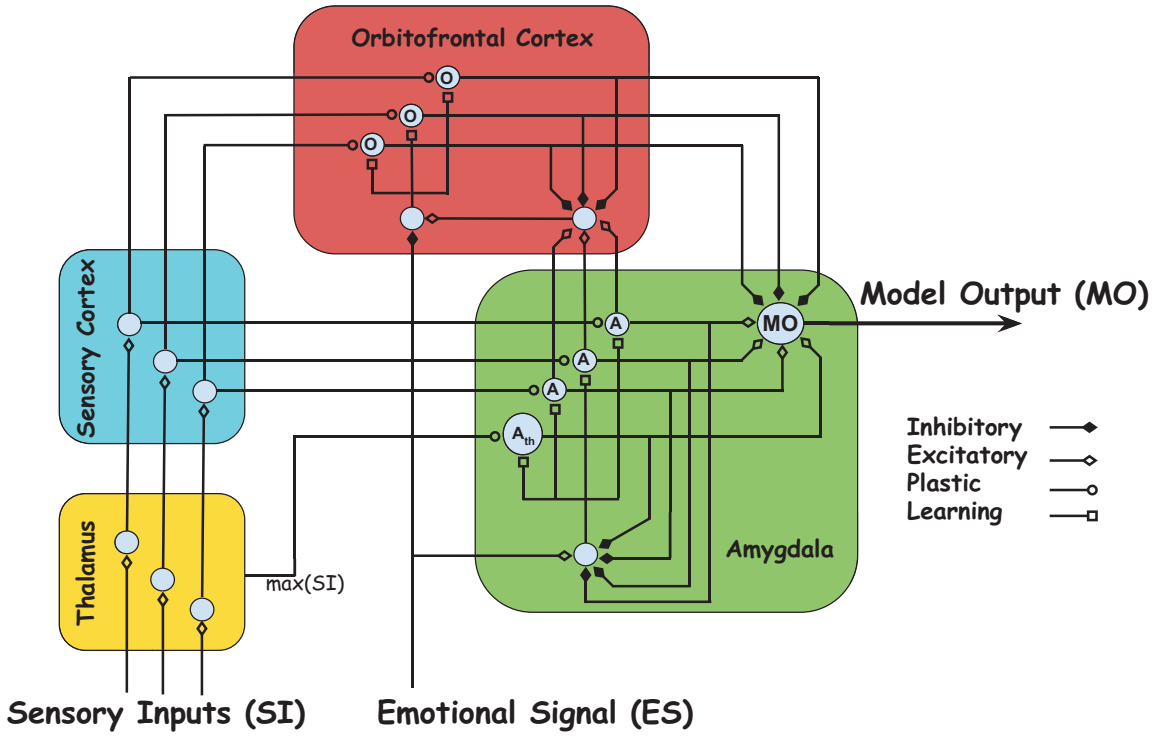

Figure 3.1: Block diagram of emotional learning.

3.2.2 Brain Emotional Learning-Based Intelligent Controller: A novel reinforcement learning approach

BELBIC was first introduced in [41], and since then, it has been successfully utilized in multiple areas. For instance, the authors in [84] applied BELBIC controller to the neuro-fuzzy model of a micro-heat exchanger. BELBIC based motion control of omnidirectional three-wheeled robots has been addressed in [85]. BELBIC was also successfully applied for the attitude control of a quad rotorcraft [43]. More recently, the navigation of an Unmanned Ground Vehicle (UGV) by employing BELBIC, was studied in [74]. Closely related, the optimal bi-objective structure emotional learning for Dynamic Voltage Restorer is addressed in [75].

Additionally, diverse research works have been devoted to the real-time im-

plementation of BELBIC. For instance, in [86], an implementation on BELBIC on a DSP was utilized for controlling an interior permanent magnet synchronous motor drive, whereas in [87] BELBIC was implemented on a Field-Programmable Gate Array for controlling an overhead traveling crane. Also, BELBIC was practically applied to the speed control of a digital servo system [46].

The capability of learning is a pivotal characteristic of an intelligent system which distinguishes it from a traditional system [41]. A common property of the learning methodologies, to cope with the changing environment in a superior way, is based on adapting the parameters of the system with various levels of computational complexity. Furthermore, it is essential for any learning methodology to have an appraisal mechanism for determining the operational condition of the system. One category of appraisal mechanisms is based on the emotional cues [40,78]. This evaluation mechanism determines how external stimuli could affect the capability of the system to effectively function in short-term and also to retain its long-term prospects for survival. *Emotional Learning*, is a learning technique that is based on emotional appraisals.

The BELBIC methodology is based on a novel architecture of the emotional learning observed in the mammalian limbic system which was proposed in [78], and is graphically represented as in Figure 3.1. This model has two main parts: *Amygdala* and *Orbitofrontal Cortex*, which are responsible for immediate learning and for inhibiting any inappropriate learning happening in the Amygdala, respectively.

Furthermore, there are two important inputs to the BELBIC model, i.e., Sensory Inputs (*SI*) and Emotional Signal (*ES*). Then, a model output equation is defined

as

$$MO = \sum_{l} A_l - \sum_{l} OC_l \tag{3.11}$$

which consists of the subtraction of Amygdala outputs (A_l) and Orbitofrontal Cortex outputs (OC_l). Here, l is the number of sensory inputs.

Amygdala and Orbitofrontal Cortex outputs are computed by the summation of all their corresponding nodes where the outputs of each node (i.e., Amygdala node (3.12) and Orbitofrontal Cortex node (3.13)) can be obtained as

$$A_l = V_l SI_l \tag{3.12}$$

$$OC_l = W_l SI_l \tag{3.13}$$

where V_l and W_l are the weights of Amygdala and Orbitofrontal Cortex, respectively, and SI_l is the l_{th} sensory input. To update V_l and W_l the following equations are used

$$\Delta V_l = K_v SI_l \max\left(0, ES - \sum_{l} A_l\right) \tag{3.14}$$

$$\Delta W_l = K_w SI_l \left(MO - ES\right) \tag{3.15}$$

where K_v and K_w are the learning rates.

Remark 2. *Assigning different values to K_v and K_w will affect the convergence performance of the algorithm (i.e., convergence of the weights of Amygdala (V_l) and Orbitofrontal Cortex (W_l)). In general, assigning the bigger values to these parameters will increase the learning rate and assigning the smaller values will decrease the learning rate. However, it is necessary that the designer select the values of these parameters, i.e., K_v and K_w, within the suitable region given in Theorem 3.*

There is an additional input (i.e., A_{th}) which directly comes from the Thalamus to the Amygdala. This input is the maximum of all SI and is obtained as

$$A_{th} = V_{th} \max(SI_l) \tag{3.16}$$

Here V_{th} is the weight and the update law is similar to (3.14).

Several methods have been developed for tuning the BELBIC parameters. For example, a particle swarm optimization-based approach was used in [79]. A Lyapunov-based tuning algorithm for a group of linear systems was adopted in [80]. Fuzzy tuning of BELBIC has been proposed and successfully utilized for controlling a robotic arm in [81]. Also, the clonal selection algorithm has been recently employed to obtain BELBIC parameters and has been applied for controlling a single link flexible joint manipulator [47]. Additionally, trial and error tuning has shown to be an effective method [43].

Although these approaches may lead to finding the optimal value for BELBIC parameters, almost most of these update laws add more computational complexity to the system that limited the practical usage of most of these algorithms. In this paper, a heuristic method was adopted for obtaining the BELBIC parameters. This approach could significantly reduce the computational complexity of defining these parameters.

3.2.3 Flocking Control

Based on the flocking problem formulated in Section 3.2.1, and by employing the BELBIC controller introduced in Section 3.2.2, the objective is to design a control signal u_i for each agent i, in such a way that the motion of all agents in the flock

represents an emergent behavior arising from simple rules that are followed by individuals, and does not involve any central coordination. Moreover, the proposed controller should keep a low level of complexity, in order to be practically implementable in the real-time flocking of multiple unmanned aircraft systems (UAS) platforms. In summary, the agents must track a virtual leader (i.e., γ-agent), while avoiding collisions with every other agent in the group, and while negotiating and avoiding obstacles encountered in the dynamic environment. The proposed control law must be capable of satisfying multiple designer objectives, specifically (i) optimization of the control effort and energy expense, (ii) robustness against environmental noise/disturbances, (iii) and capability of handling the model uncertainty. These objectives must be accomplished without increasing the complexity of the system.

3.3 Flocking Control of MAS using BELBIC

3.3.1 System Design

Generally, there are two methodologies, the so-called direct and the indirect method [41], in the utilization of cognitive and/or intelligent control. In the direct method, the intelligent and/or cognitive model is employed as a controller block, while in the indirect method, the intelligent and/or cognitive model is used for obtaining the controller's parameters. In our work, the first method i.e., direct approach has been adopted, where a biologically-inspired intelligent methodology based on the novel architecture of emotional learning in mammal's brain is employed as the controller block.

In this regard, we need to figure out how to embed this model as a controller within the overall system architecture. In particular, there is no unique scheme for utilizing BELBIC in the system architecture. It is based on a fact that the pivotal property of the BELBIC model is its flexibility in attaining multiple objectives by assigning the different emotional signals and sensory inputs.

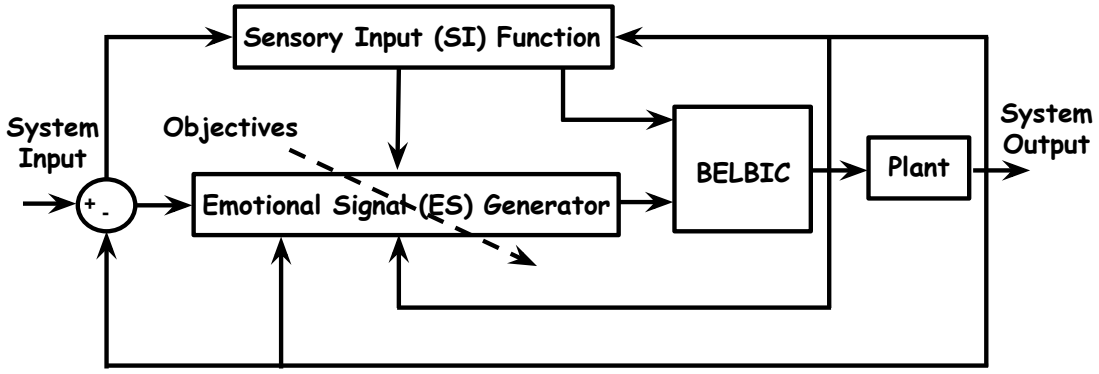

Figure 3.2: BELBIC in the control loop.

The BELBIC architecture implemented in this work is shown in Figure 3.2. It is a closed loop configuration which consists of BELBIC, Emotional signal generator (ES), Sensory inputs function (SI), and the plant. This architecture implicitly demonstrates the overall emotional based control concept which consists of the critic, learning algorithm, and the action selection mechanism [41].

In this context, the given controller acts as an intelligent distributed flocking controller. It should be noted that this structure is utilized for each individual agent separately.

The details of the learning algorithm are given in the following sections.

3.3.2 Emotional Signal and Sensory Input Development

From a biological point of view, the emotional signal is a generic signal which is generated internally and could represent different reinforcing inputs from Thalamus, Hypothalamus, and parts of the Basal Ganglia [40,78]. Artificially simulating the model, different parts of the system could generate the emotional signal to reflect any related criteria.

The primary idea behind the emotional learning based control and/or decision-making in this work is to produce the action (i.e., output) that regulates the emotional signal (i.e., maximizes the emotional reward or minimizes the emotional stress), while different sets of sensory inputs are received by the system [41]. The current situation of the system is represented by the received sensory inputs and the emotional signals represent how satisfactory the performance of the system is. In other words, the emotional signal represents the condition of the system by considering the particular objective of interest.

Essentially, BELBIC is producing its actions based on the Emotional signal (ES) and Sensory input (SI). In general, SI and ES are defined as

$$SI_i = G_i\left(y, e, u, r\right) \tag{3.17}$$

$$ES_i = F_i\left(y, e, u, r\right) \tag{3.18}$$

where y, e, u, and r are system output, system error, control effort, and system input, respectively. The designer can implicitly decide the control objectives by choosing the adequate ES. For example, it is possible to choose the ES to reduce the energy expense, to maintain network connectivity, or to achieve better target tracking performance, etc.

Aiming at designing a practical controller appropriate for implementation in real MAS, the BELBIC design proposed here will focus on keeping small the control effort and the energy of the flock, in order to avoid actuator saturation, while also considering the perfect flocking behavior, which in turn will extend the energy available for performing the mission. To fulfill this objective, the ES is designed in such a way that the increase in control effort will represent a negative emotion, e.g., stress, which is then taken as an evidence that the system performance is not satisfactory which leads to control effort reduction and energy efficiency. Then, for each agent i, the SI_i and ES_i are designed as

$$SI_i = K^{\alpha}_{i,SI} u^{\alpha}_i + K^{\beta}_{i,SI} u^{\beta}_i + K^{\gamma}_{i,SI} u^{\gamma}_i \tag{3.19}$$

$$ES_i = K^{\alpha}_{i,ES} u^{\alpha}_i + K^{\beta}_{i,ES} u^{\beta}_i + K^{\gamma}_{i,ES} u^{\gamma}_i \tag{3.20}$$

where $K^{\alpha}_{i,SI}$, $K^{\beta}_{i,SI}$, $K^{\gamma}_{i,SI}$, $K^{\alpha}_{i,ES}$, $K^{\beta}_{i,ES}$, $K^{\gamma}_{i,ES}$ are positive gains. By assigning different values to these gains, the ES will change its influence on the system behavior. In this work, identical values are used since the objective is to reduce the overall control effort and energy of the system.

Remark 3. *Assigning different values to parameters in SI_i and ES_i (i.e., $K^{\alpha}_{i,SI}$, $K^{\beta}_{i,SI}$, $K^{\gamma}_{i,SI}$, $K^{\alpha}_{i,ES}$, $K^{\beta}_{i,ES}$, $K^{\gamma}_{i,ES}$) would have different impacts on the performance of the whole system. In other words, these gains provide the design freedom allowing different priorities for each one of the three different components of flocking. A high gain will make a specific component more important than the others. However, since our objective in this work is to reduce the control effort and the energy of the flock, while also considering the perfect flocking behavior, our only design constraint in choosing these values is to avoid providing any privileges to a specific component of the flocking elements (i.e., Collision Avoidance, Obstacle Avoidance, and Navigational Feedback).*

One of the main advantages of designing the ES as in equation (3.20) is that the flocking objective can be adaptive according to the scenario faced by the system. For example, the pseudo-code in **Algorithm** 2 defines the gain $K^{\beta}_{i,ES}$ depending on the situation faced by the flock: the presence of obstacles, or the obstacle-free environment.

In other words, when no obstacle is within a sensing range of an agent, the obstacle avoidance part will be canceled out from the emotional signal. If the agent senses an obstacle, depending on the size of the obstacle, the distance of obstacle with respect to the agent, etc., a value will be assigned to maintain the impact of the obstacle avoidance part of the emotional signal for the corresponding agent.

Algorithm 2 : Adaptive gain $K^{\beta}_{i,ES}$.

if agent i senses an obstacle **then**
  set $K^{\beta}_{i,ES} =$ computed value with respect to obstacle
else
  set $K^{\beta}_{i,ES} = 0$
end if

Designing the appropriate functions for SI_i and ES_i, our goal is to compute the model output (MO_i) in the following section. Then, the model output will be employed as our intelligent controller output for flocking control of MAS.

3.3.3 Learning-based Flocking Control

Flocking is a very nice case for employing learning-based multi-objective approaches like BELBIC, since it consists of multiple performance considerations to be taken into account all at the same time. Learning the flocking behavior both in presence of obstacles and/or in obstacle free environment while considering the problem

of handling the system uncertainties and disturbances and also being appropriate for real-time implementation due to its low computational complexity and control effort reduction capability, leads us to take advantage of the computational model of emotional learning in the mammalian limbic system, i.e., BELBIC.

From equations (3.19)-(3.20), the BELBIC for flocking of MAS is defined as

$$\begin{aligned} u_i^{BEL} &= \sum_i V_i.SI_i - \sum_i W_i.SI_i \\ &= \sum_i V_i.\left(K_{i,SI}^{\alpha}.u_i^{\alpha} + K_{i,SI}^{\beta}.u_i^{\beta} + K_{i,SI}^{\gamma}.u_i^{\gamma}\right) \\ &\quad - \sum_i W_i.\left(K_{i,SI}^{\alpha}.u_i^{\alpha} + K_{i,SI}^{\beta}.u_i^{\beta} + K_{i,SI}^{\gamma}.u_i^{\gamma}\right) \end{aligned} \tag{3.21}$$

Here $i = 1, ..., n$ and n is the number of agents. By considering the results obtained from Theorem 3 and by substituting the Emotional Signal with equation (3.20) the BEL-Flocking model output for MAS could be obtained as follows:

$$MO_i = ES_i = K_{i,ES}^{\alpha}u_i^{\alpha} + K_{i,ES}^{\beta}u_i^{\beta} + K_{i,ES}^{\gamma}u_i^{\gamma} \tag{3.22}$$

which is exactly demonstrating the flocking behavior of the proposed method.

The overall BELBIC-based flocking methodology proposed is summarized as pseudo-code in **Algorithm** 3.

Stability Analysis

The convergence of the weights of Amygdala (V_i) and Orbitofrontal Cortex (W_i) are presented in Theorem 3.

Algorithm 3 : The BELBIC-based intelligent flocking control for MAS.

Initialization:
Set $V_i = 0$, $W_i = 0$, and $V_{th} = 0$, for $i = 1, ..., n$.
Define $ES_i =$ Objective function, for $i = 1, ..., n$.

for each iteration $t = t_s$ **do**
  for each agent i **do**
    Compute $SI_i = K^{\alpha}_{i,SI} u^{\alpha}_i + K^{\beta}_{i,SI} u^{\beta}_i + K^{\gamma}_{i,SI} u^{\gamma}_i$
    Define $K^{\beta}_{i,ES}$ using **Algorithm** 2
    Compute $ES_i = K^{\alpha}_{i,ES} u^{\alpha}_i + K^{\beta}_{i,ES} u^{\beta}_i + K^{\gamma}_{i,ES} u^{\gamma}_i$
    Compute $A_i = V_i SI_i$
    Compute $OC_i = W_i SI_i$
    Compute $A_{th} = V_{th} \max(SI_i)$
    Compute $MO_i = \sum_i A_i - \sum_i OC_i$
    Update V_i
    Update W_i
    Update V_{th}
  end for
end for

Theorem 3. *Given the BELBIC design as* (3.19)–(3.22)*, there exists the positive BELBIC tuning parameter, K_v, K_w satisfying*

I. $\left|\left[1 - K_v (SI_i)^2\right]\right| < 1$

II. $\left|\left[1 - K_w (SI_i)^2\right]\right| < 1$

such that the MAS's estimated weights of Amygdala (V_i) and Orbitofrontal Cortex (W_i) converge to desired targets asymptotically.

Proof. See **Appendix A** □

Remark 4. *Based on the BELBIC theory [41] and* (3.21)*, the optimal flocking control of MAS can be obtained while the estimated weights of Amygdala (V_i) and Orbitofrontal Cortex (W_i) are converging to desired targets. According to Theorem 3, estimated weights converge to desired targets asymptotically. Therefore, the designed BELBIC input U_i^{BEL}* (3.21) *converges to the optimal flocking control of MAS asymptotically.*

3.4 Simulation Results

This section presents simulation results demonstrating the performance of the BELBIC for flocking of MAS under four different scenarios:

i. 2-dimensional obstacle-free environment (Subsection 3.4.1)
ii. 2-dimensional environment in presence of obstacles (Subsection 3.4.2)
iii. 3-dimensional obstacle-free environments (Subsection 3.4.3)
iv. 3-dimensional environment in presence of obstacles (Subsection 3.4.4)

Since most of the flocking algorithms in the literature have investigated 2D agents scenarios, in order to be able to evaluate our proposed methodology in comparison with them, we present the 2D scenarios in subsections 3.4.1 and 3.4.2 for agents operating in the obstacle-free environment and in presence of obstacles, respectively. The 3D scenarios presented in 3.4.3 and 3.4.4 are included because our ultimate goal in this work is to implement our proposed method in a real-time application considering multiple UAS, therefore, we need to evaluate the proposed approach in a 3-dimensional environment in order to be able to experimentally apply it in a real-world application. We provided the comparison of the controller

signals of the proposed method with respect to other methods in order to validate the performance of the proposed controller in terms of reducing the controller effort and the kinetic energy of the system.

In both 2-dimensional scenarios, a total of 150 agents where employed, while a total of 50 agents were employed in both 3-dimensional scenarios. Initial velocities in all cases are equal to zero, and positions randomly distributed in a squared area. The following parameters are used through all the simulations: $r = 1.2d_\alpha$, $d^{'} = 0.6d_\alpha$, $r^{'} = 1.2d^{'}$ and in the 2-dimensional cases $d_\alpha = 7$, while $d_\alpha = 4$ for the 3-dimensional cases. For the σ-norm the parameter $\epsilon = 0.1$, for $\phi(z)$ the parameters $a = b = 5$, for the bump function $\phi_\alpha(z)$ we use $h = 0.2$, and for $\phi_\beta(z)$ we use $h = 0.9$. The remaining parameters of the algorithm are specified in each subsection, as needed.

3.4.1 BELBIC for Flocking in 2-Dimensional Obstacles-Free Environment

Figure 3.3 and Figure 3.4 show two snapshots of the simulation in the obstacle-free environment. Figure 3.3 shows the 150 agents in their initial positions at $t = 0s$. It can be seen that the randomly distributed agents start forming an α-lattice. Figure 3.4 shows the agents at $t = 70s$ where they are flocking and have successfully formed a connected network.

For comparison purposes, two similar experiments were performed, but using the flocking algorithms introduced in [1] and Multirobot Cooperative Learning for Predator Avoidance (MCLPA) flocking algorithm introduced in [2] instead of

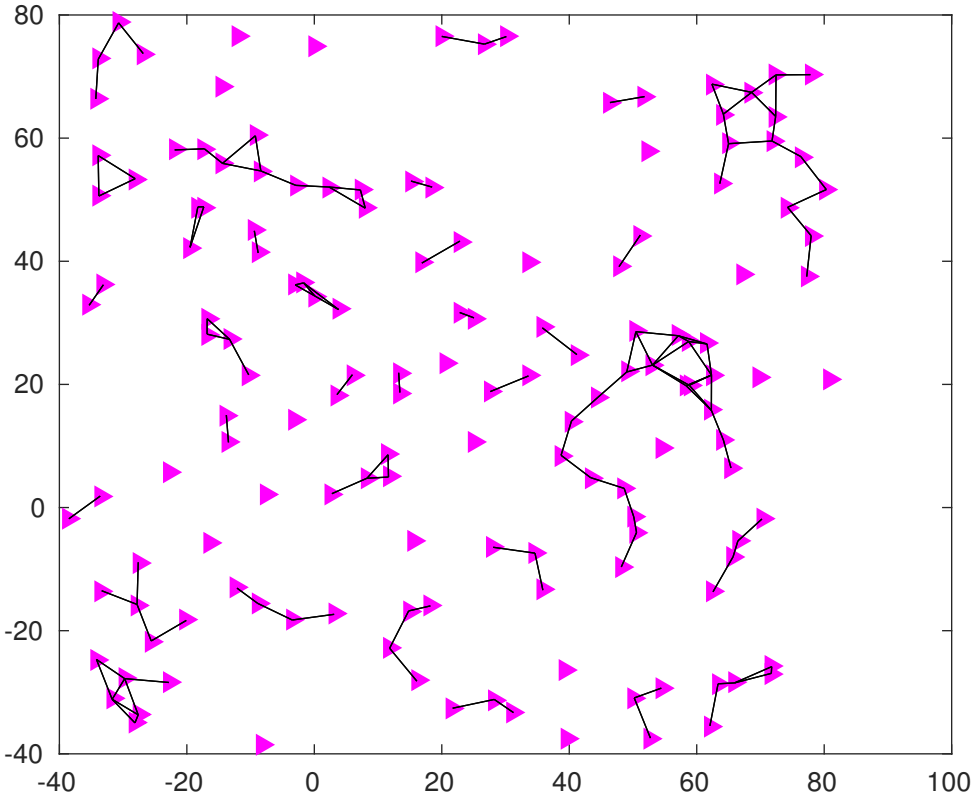

Figure 3.3: BELBIC-based Flocking of MAS. Simulation in an obstacle-free environment. 150 agents randomly distributed in a squared area at $t = 0s$

.

the BELBIC-based flocking. Figure 3.5, Figure 3.6, and Figure 3.7 show the control signals generated by the conventional flocking method in [1], MCLPA flocking method in [2], and the BELBIC-based flocking, respectively. The Table 3.1 presents some characteristics of the control signals generated by all three conventional, MCLPA, and the BELBIC-based flocking strategies in an obstacle-free environment.

Figure 3.8 shows the Mean Square Value $\frac{1}{n}\sum_{i=1}^{n}(u_i)^2$ of the control effort generated by the overall group of agents, for the flocking methods in [1], the MCLPA flocking strategy in [2], and the BELBIC-based flocking in obstacle-free environ-

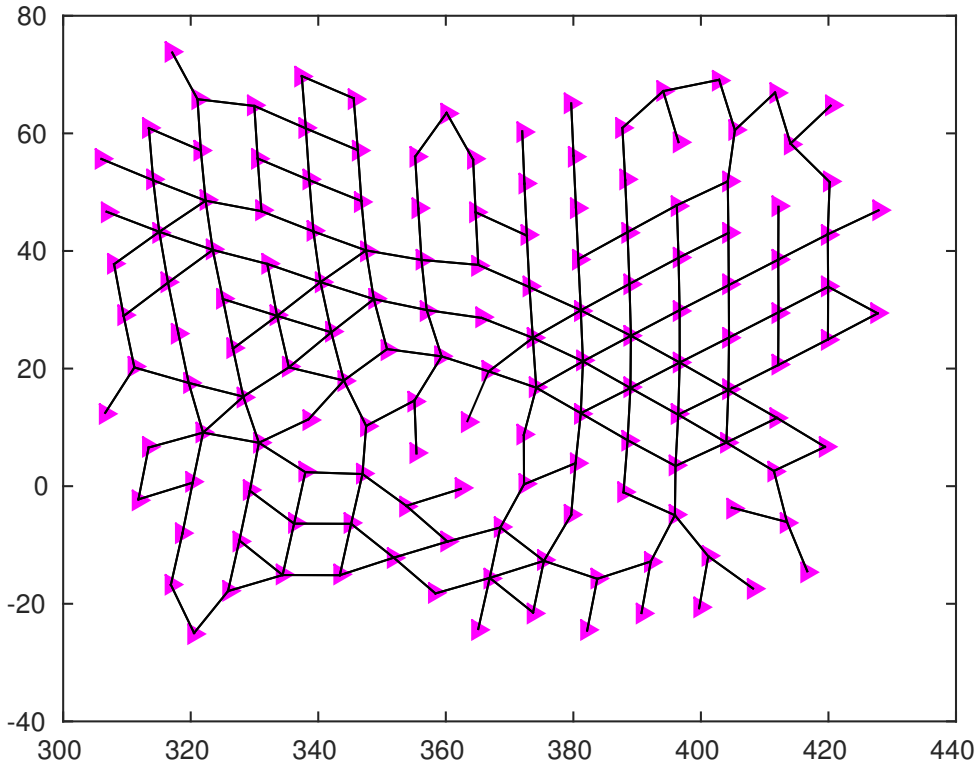

Figure 3.4: BELBIC-based Flocking of MAS. Simulation in an obstacle-free environment. At $t = 70s$ the 150 agents are flocking and have successfully formed a connected network.

Table 3.1: Characteristics of the control signals generated by all three flocking strategies in an obstacle free environment.

	Flocking in [1]	MCLPA [2]	BELBIC-based
Max Value	37.8306	43.9181	6.0529
Min Value	-5.8222	-7.1224	-1.2349
Mean Value	0.1343	0.1431	0.1334
Standard Deviation	$1.0979E-04$	$1.3051E-04$	$0.6905E-04$

ment. The plot shows that, despite all methods are able to accomplish flocking of MAS, the control signals generated by the BELBIC-based flocking are smaller, and therefore more appropriate to implement in real-robots.

Figure 3.9 presents the Kinetic Energy $K(v) = \frac{1}{2}\sum_i \| v_i\|^2$ associated to the

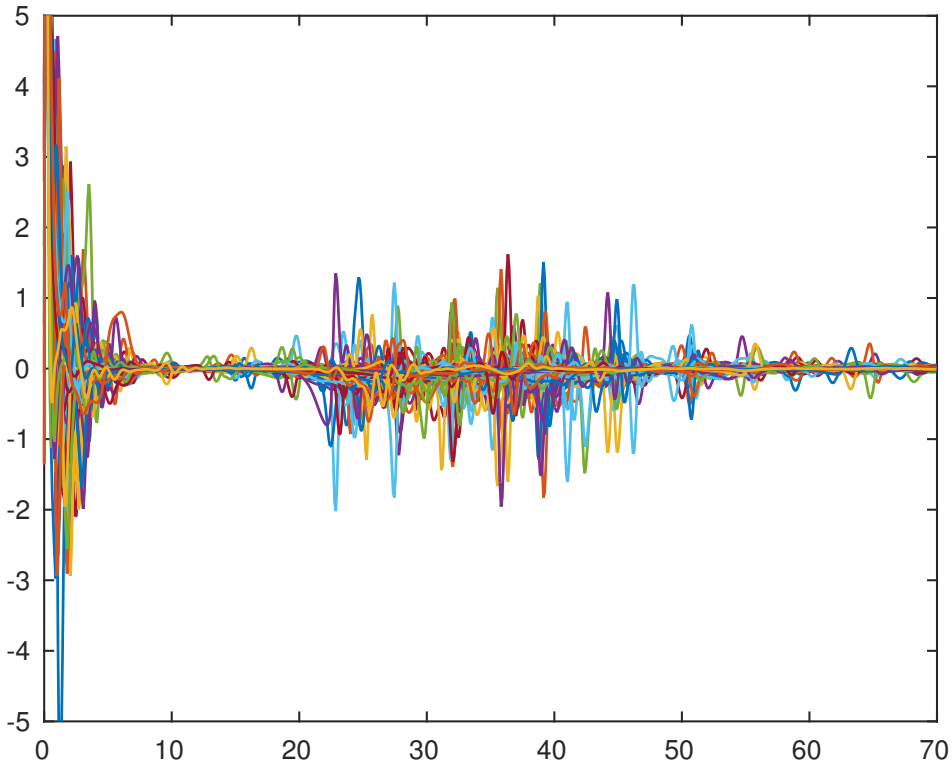

Figure 3.5: Control signals generated by the flocking algorithm proposed in [1] in an obstacle-free environment.

overall group of agents, for the flocking methods in [1], the MCLPA flocking strategy in [2], and the BELBIC-based flocking, in an obstacle free environment. Here $v_i = p_i - \bar{p}$ and $\bar{p} = \frac{1}{n}\sum_i p_i$. The plot shows that the Kinetic Energy associated to the BELBIC-based flocking is smaller, and therefore more appropriate for real-time implementations.

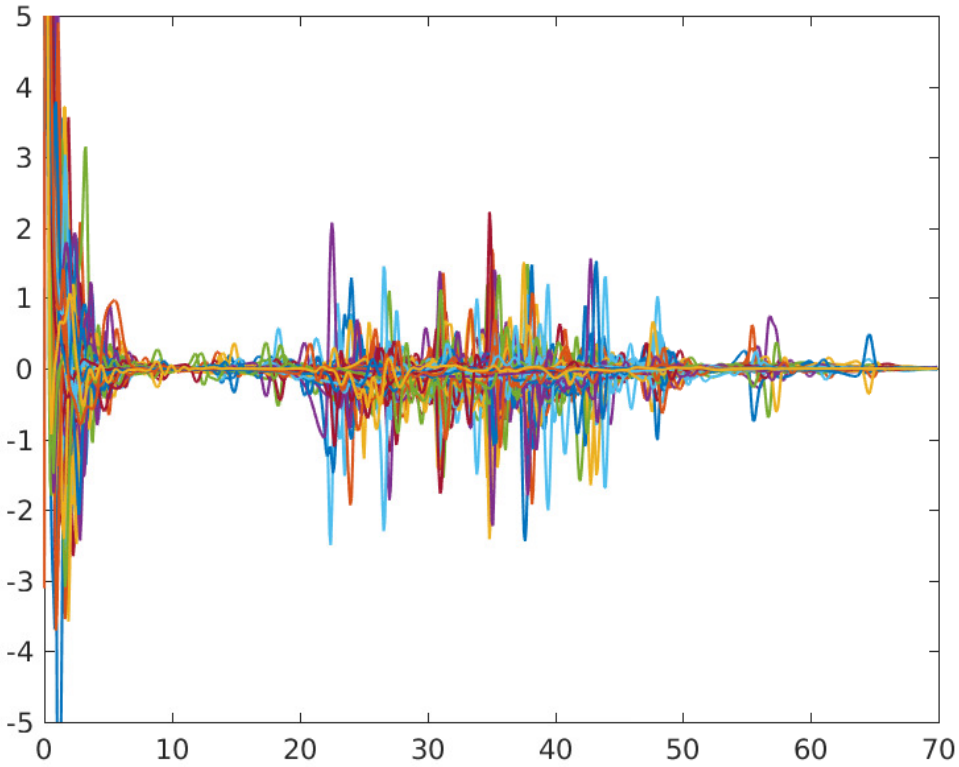

Figure 3.6: Control signals generated by the MCLPA flocking algorithm proposed in [2] in an obstacle-free environment.

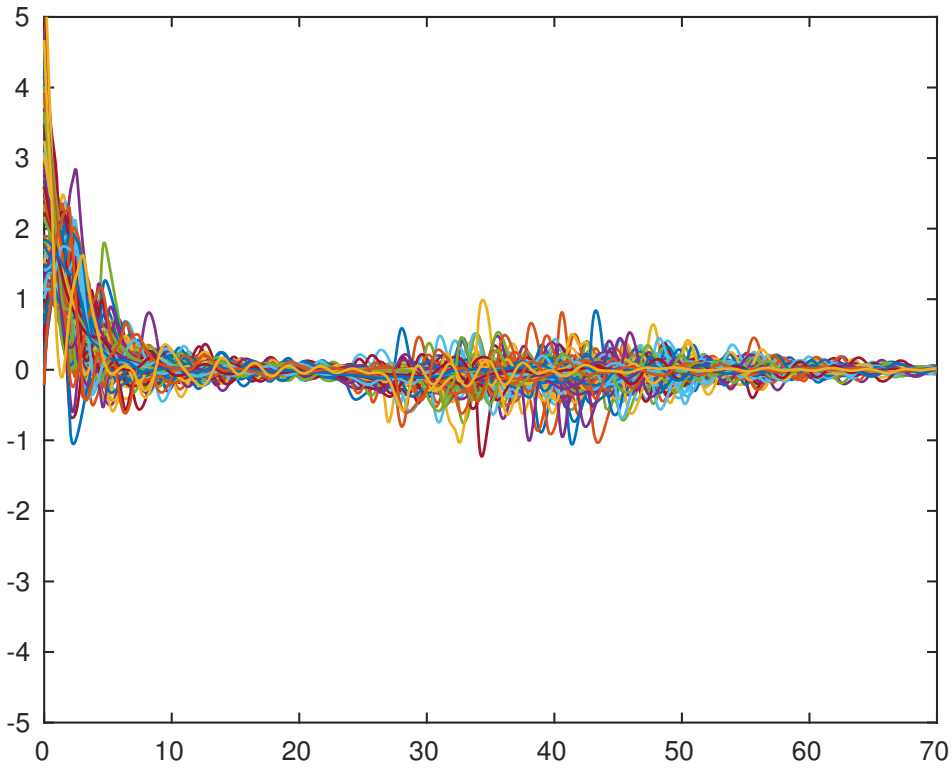

Figure 3.7: Control signals generated by the BELBIC-based flocking algorithm in an obstacle-free environment.

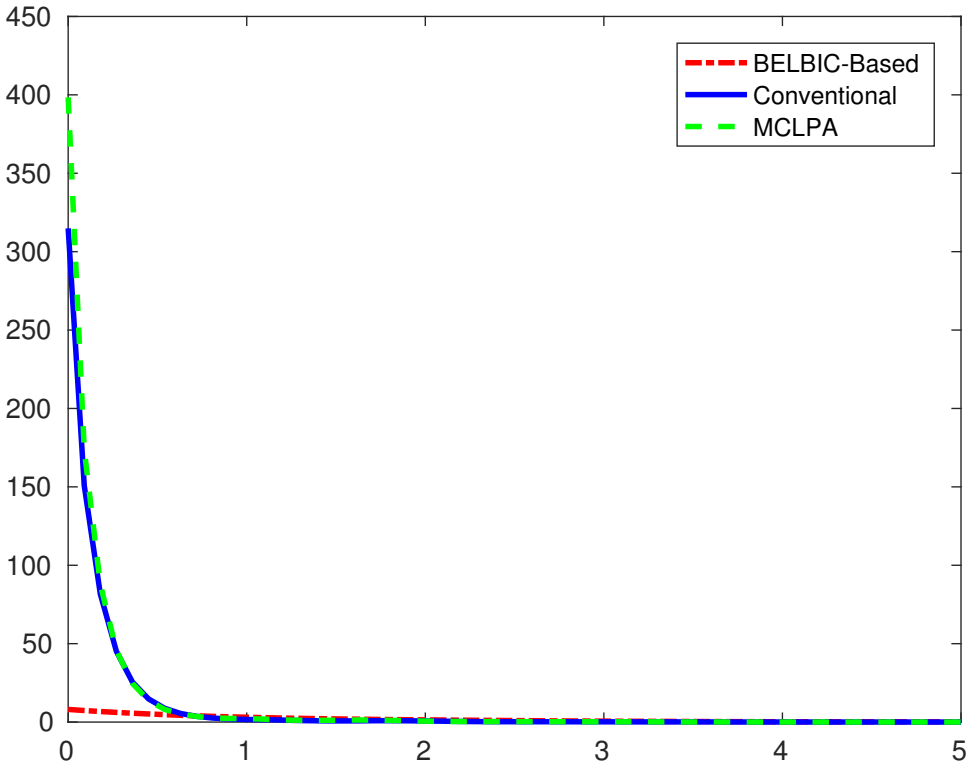

Figure 3.8: Mean Square value of all agents control effort in obstacle-free environment. The BELBIC-based flocking is presented in dot-dashed red, the MCLPA flocking in [2] in dashed green, and the flocking in [1] in solid blue. Notice that the control signals generated by the BELBIC-based flocking are smaller, and therefore more appropriate to implement in real-robots.

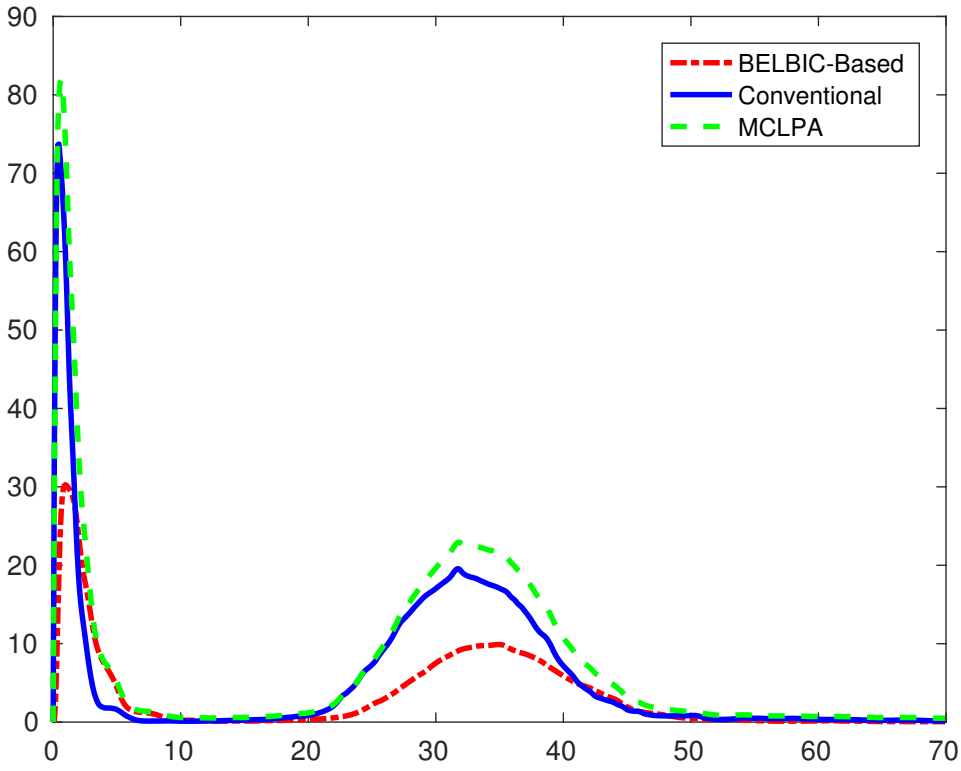

Figure 3.9: Kinetic Energy of all agents in an environment free of obstacles. The BELBIC-Based flocking energy is shown as a dot-dashed red line, while the energy of the MCLPA flocking in [2] in dashed green and the energy of the flocking in [1] is shown as a solid blue line. The Kinetic Energy associated to the BELBIC-based flocking is smaller, and therefore more appropriate for real-time implementations.

3.4.2 BELBIC-based flocking in 2-Dimensional Environment with Obstacles

For this scenario, a set of obstacles were created, with the characteristics contained in the following matrix

$$OBS = \begin{bmatrix} 100 & 110 & 120 & 130 & 150 & 160 \\ 20 & 60 & 40 & -20 & 40 & 0 \\ 10 & 4 & 2 & 5 & 5 & 3 \end{bmatrix}$$

Here, the first and second row represent the horizontal and vertical coordinates of each obstacle, respectively, and the last row represents the obstacle's radius.

Figure 3.10 and Figure 3.11 show two snapshots of the simulation in presence of obstacles. Figure 3.10 shows the six obstacles, as well as the 150 agents at $t = 0s$ in their initial random positions. It can be observed that the agents start to form the α-lattice, while moving towards the area where the obstacles are located. Figure 3.11 shows the agents at $t = 20s$, where it can be observed that they managed to successfully negotiate the obstacles without any collision. The agents finally form the connected network after completely passing all the obstacles, as can be seen in Figure 3.12.

For comparison purposes, two similar experiments were performed, but using the flocking algorithms introduced in [1] and MCLPA flocking algorithm introduced in [2] instead of the BELBIC-based flocking algorithm. Figure 3.13, Figure 3.14, and Figure 3.15 show the control signals generated by the conventional flocking method in [1], MCLPA flocking method in [2], and the BELBIC-based flocking approach, respectively. From these plots, it is possible to observe that,

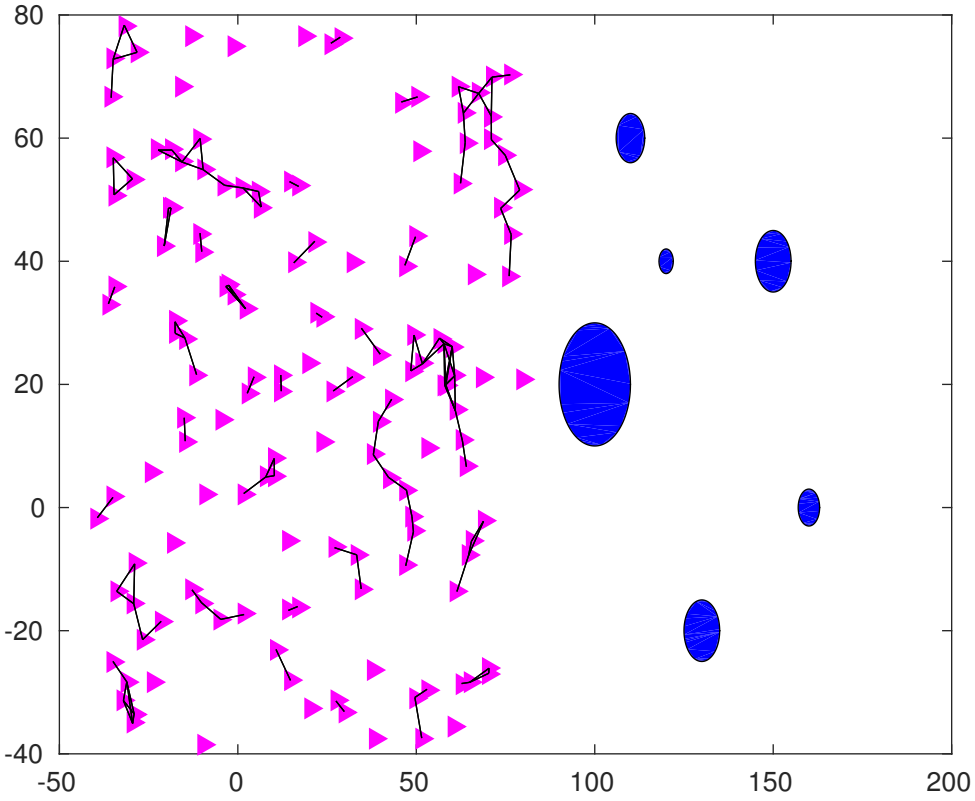

Figure 3.10: Simulation of the BELBIC-based flocking algorithm for MAS evolving in an environment with obstacles. At $t = 0s$, the 150 agents are randomly distributed. The obstacles appear as circles of different sizes.

despite all methods achieve MAS flocking, the controller effort is smaller for the BELBIC-based flocking methodology. The Table 3.2 presents some characteristics of the control signals for all three flocking methods introduced in [1], in [2], and the BELBIC-Based flocking in the presence of obstacles.

Table 3.2: Characteristics of the control signals generated by both flocking strategies in the presence of obstacles.

	Flocking in [1]	MCLPA [2]	BELBIC-based
Max Value	48.0637	65.4242	12.1297
Min Value	-86.5126	-98.3016	-21.1641
Mean Value	0.9791	1.2055	0.4306
Standard Deviation	0.0063	0.0097	0.0011

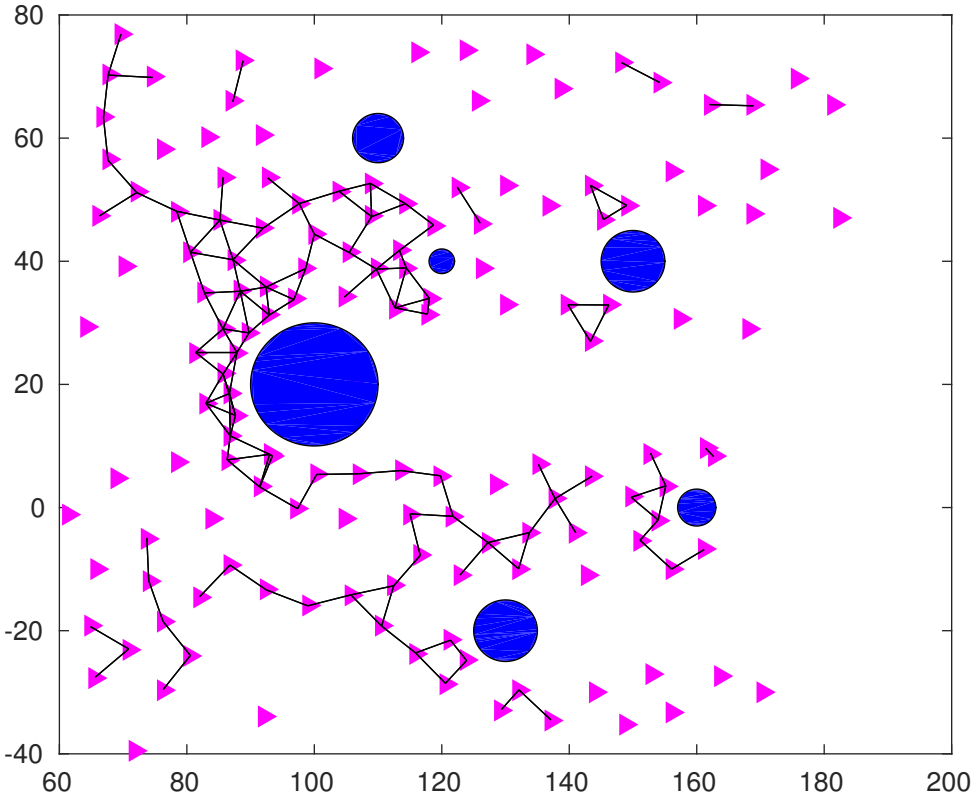

Figure 3.11: Simulation of the BELBIC-based flocking algorithm for MAS evolving in an environment with obstacles. At $t = 20s$, the 150 agents are successfully negotiating the obstacles without any collision.

Figure 3.16 show the Mean Square value $\frac{1}{n}\sum_{i=1}^{n}(u_i)^2$ of the control effort generated by the overall group of agents, for the flocking methods in [1], the MCLPA flocking strategy in [2], and the BELBIC-based flocking, when evolving in an environment with obstacles.

Figure 3.17 presents the Kinetic Energy $K(v) = \frac{1}{2}\sum_i \| v_i\|^2$ associated to the overall group of agents, for the flocking in [1], the MCLPA flocking strategy in [2], and the BELBIC-based flocking, when agents are evolving in an environment with obstacles. Here $v_i = p_i - \bar{p}$ and $\bar{p} = \frac{1}{n}\sum_i p_i$. The plot shows that the Kinetic Energy associated to the BELBIC-based flocking is smaller than the MCLPA flocking

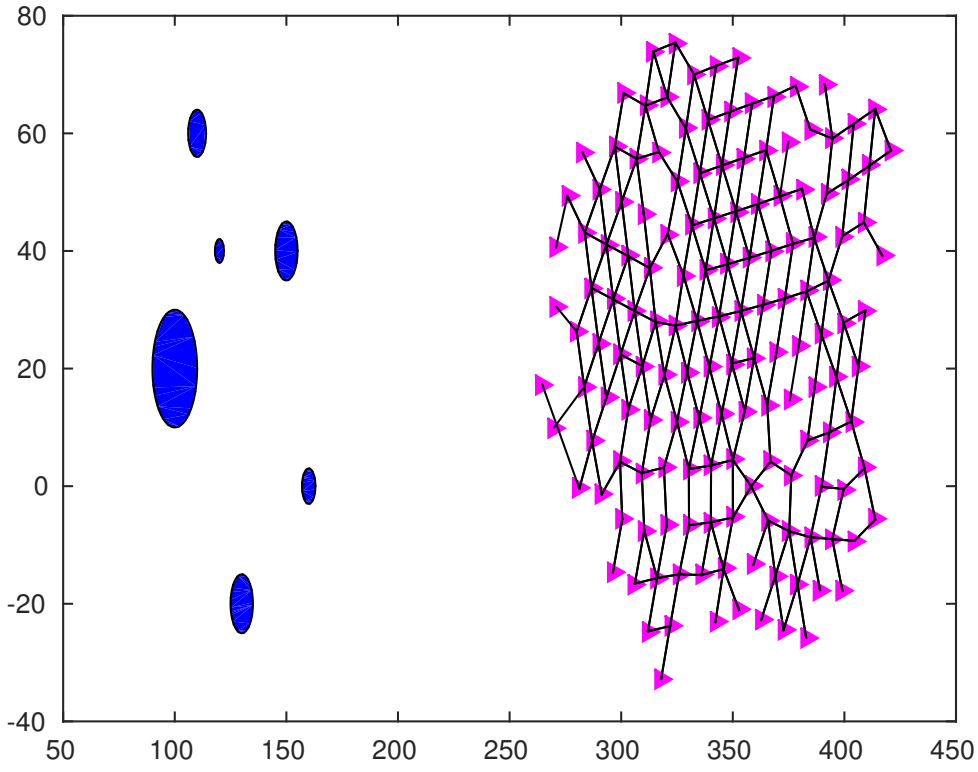

Figure 3.12: Simulation of the BELBIC-based flocking algorithm for MAS evolving in an environment with obstacles. At $t = 70s$, the 150 agents have successfully navigated through the obstacles, without any collision, and have formed a connected network.

strategy in [2] and similar to the Kinetic Energy generated by the flocking in [1]. However, it is worth noticing that the BELBIC-based flocking signal is smoother, and therefore more appropriate for real-time implementations.

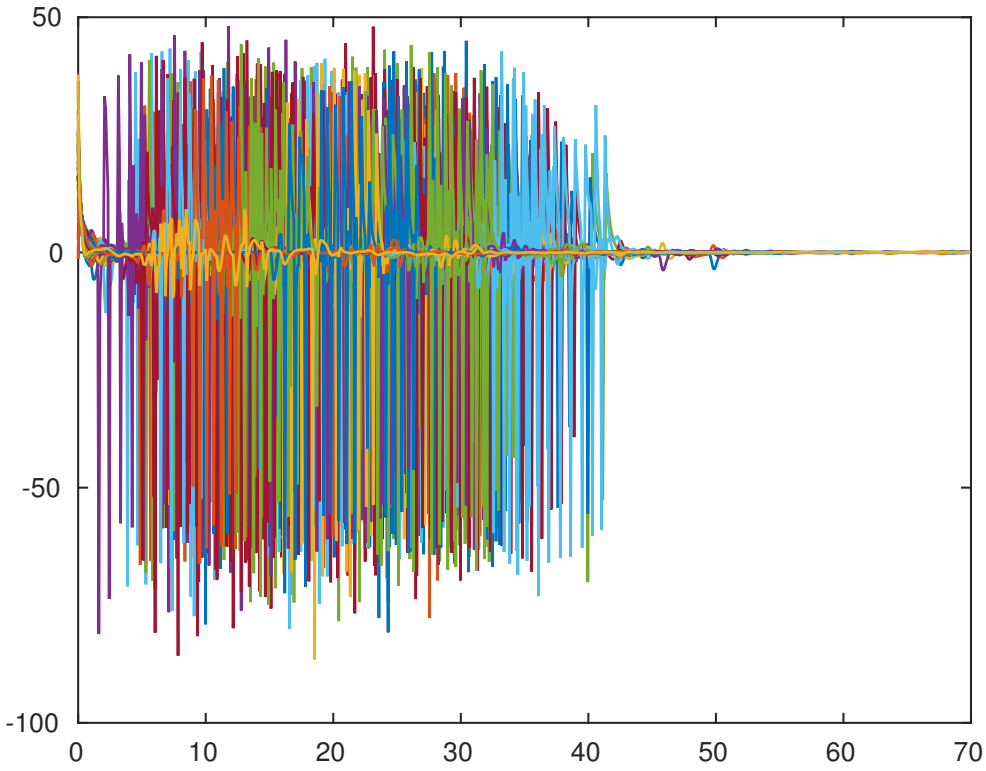

Figure 3.13: Control signals generated when using the flocking method introduced in [1], in an environment with obstacles.

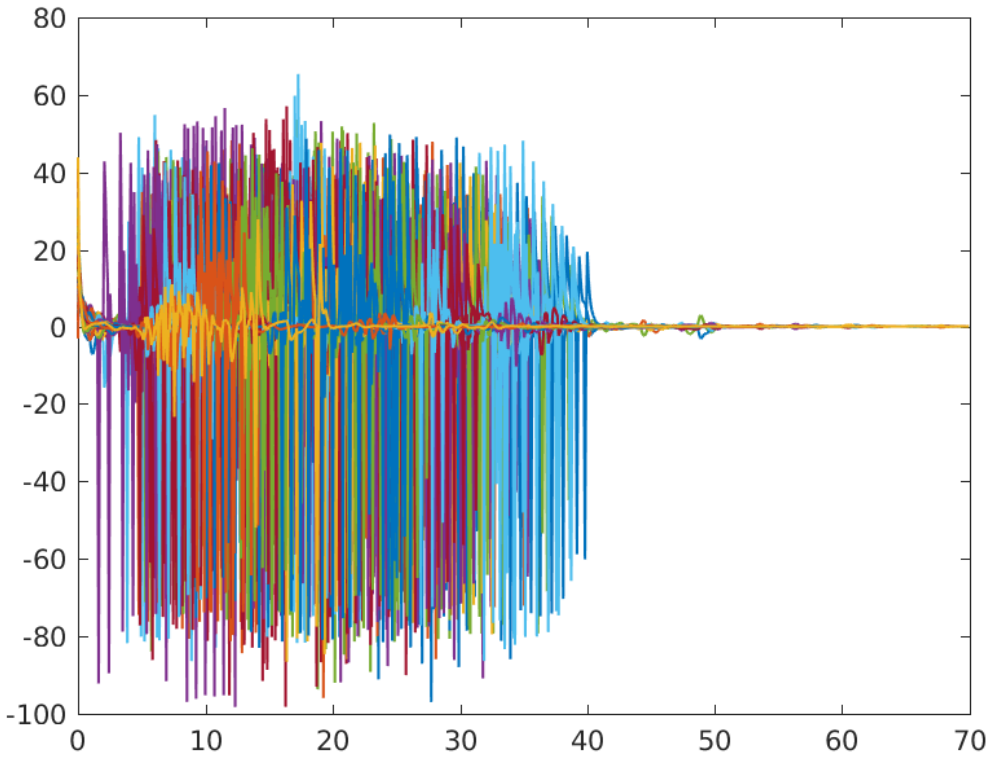

Figure 3.14: Control signals generated when using the MCLPA flocking method introduced in [2], in an environment with obstacles.

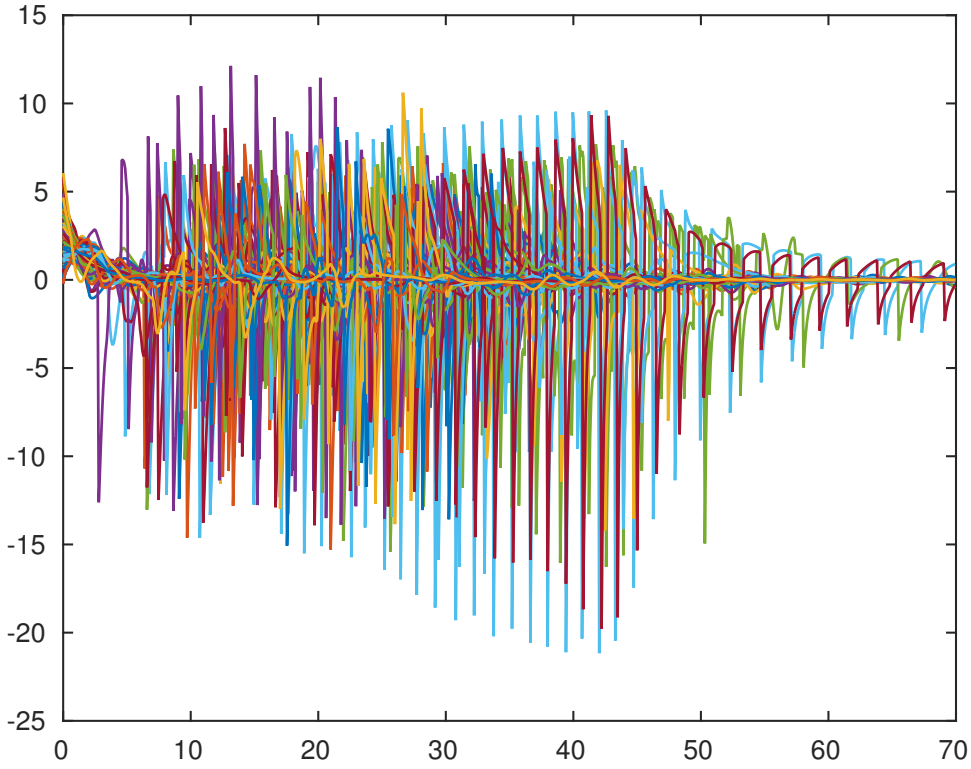

Figure 3.15: Control signals generated when using the BELBIC-based flocking method, in an environment with obstacles.

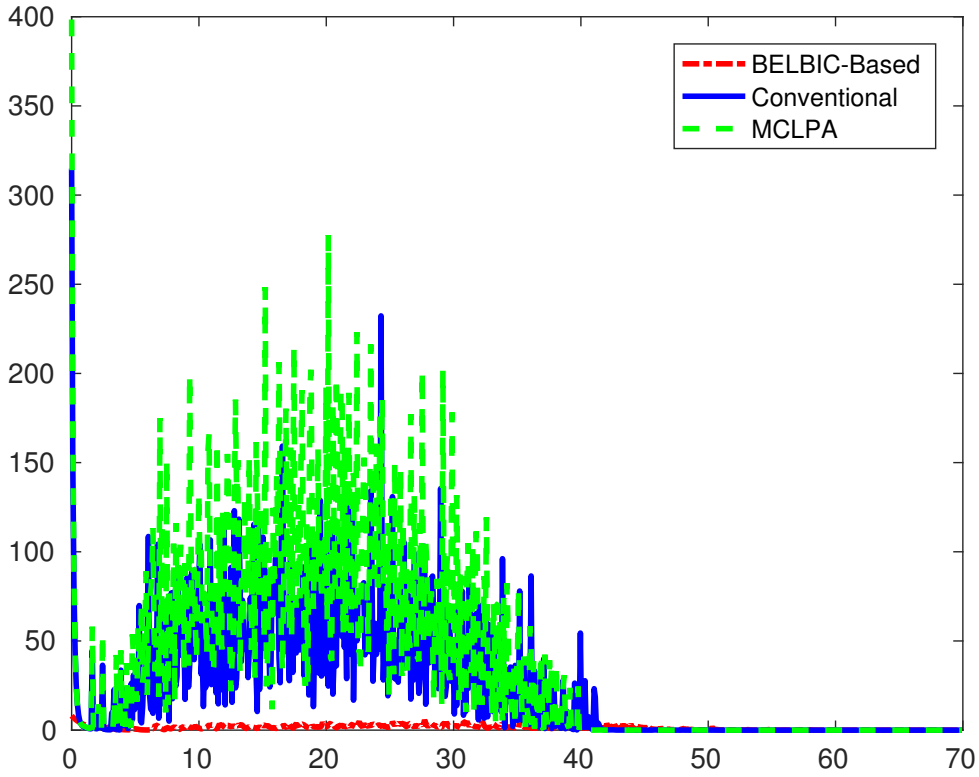

Figure 3.16: Mean Square value of the control effort generated by the overall group of agents when flocking in an environment with obstacles. The BELBIC-based flocking is presented in dot-dashed red, the MCLPA flocking strategy in [2] in dashed green, and the flocking in [1] in solid blue. Notice that the control signals generated by the BELBIC-based flocking are smaller, and therefore more appropriate to implement in real-robots.

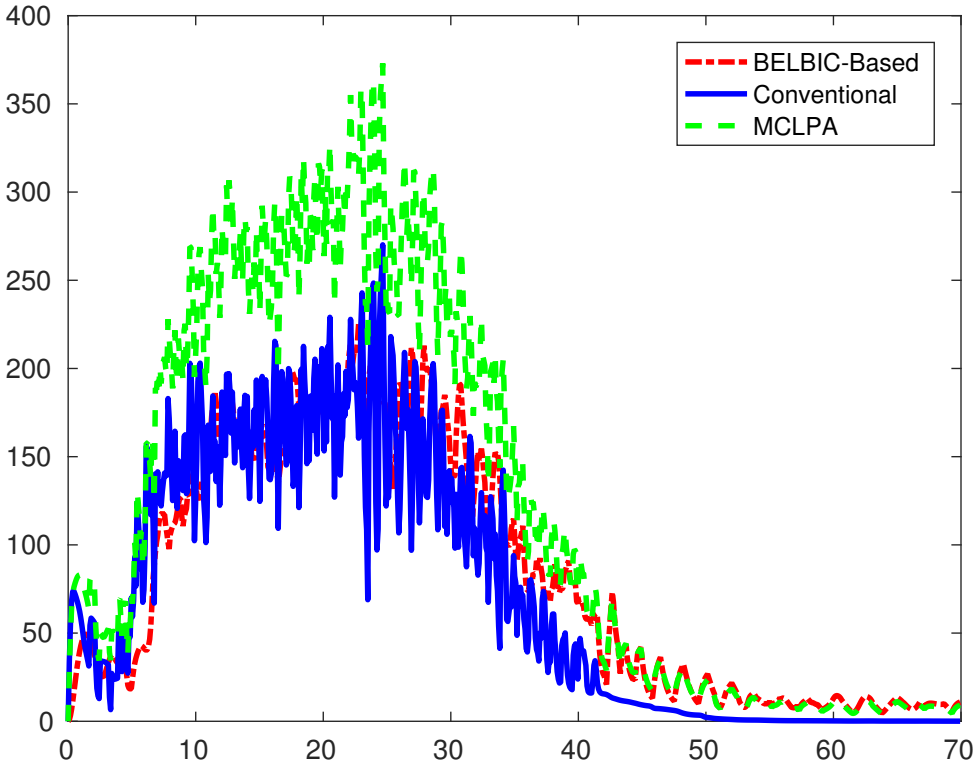

Figure 3.17: Kinetic Energy associated with the overall group of agents when evolving in an environment with obstacles. The BELBIC-Based flocking appears as a dot-dashed red line, while the MCLPA flocking strategy in [2] in dashed green, and the flocking method in [1] appears as a solid blue line. The Kinetic Energy of the BELBIC-based flocking is smaller than the MCLPA flocking strategy in [2] and similar to the Kinetic Energy generated by the flocking in [1]. However, the BELBIC-based flocking signal is smoother, and therefore more appropriate for real-time implementations.

3.4.3 Flocking in a 3-dimensional Obstacle-Free Environment

Figure 3.18 shows the 50 UAS in their initial positions, at $t = 0s$. Figure 3.19 shows the UAS at $t = 70s$ where they have successfully formed a 3D connected network.

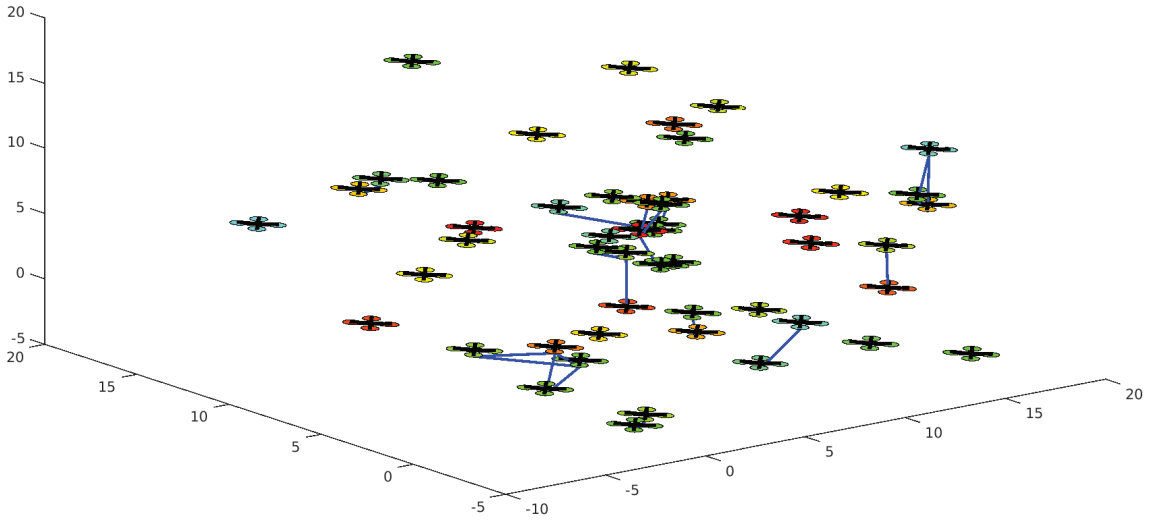

Figure 3.18: BELBIC-inspired Flocking of MAS in a 3D obstacle-free environment, at $t = 0s$.

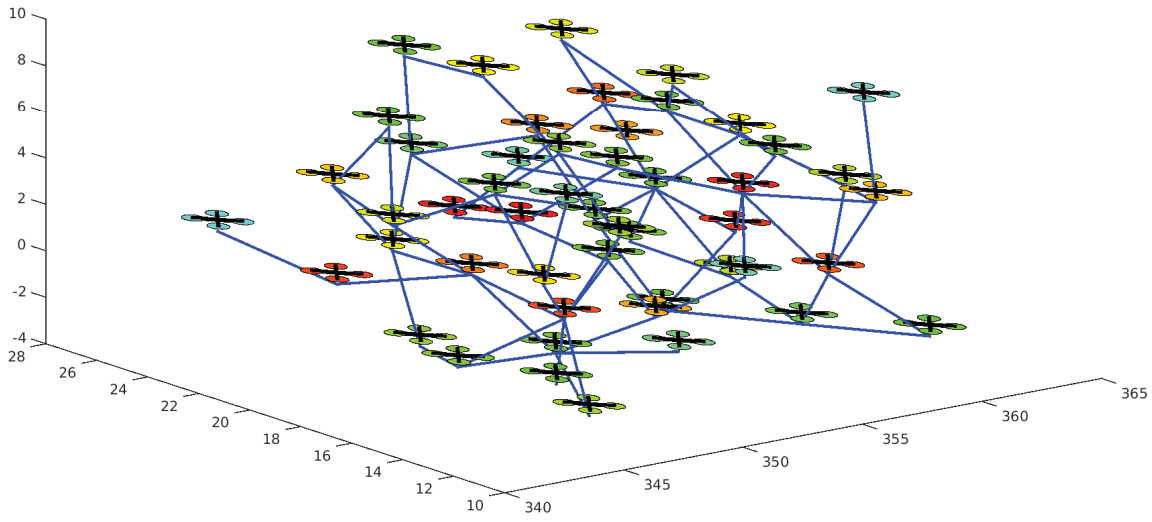

Figure 3.19: BELBIC-inspired Flocking of MAS. Simulation in a 3D obstacle-free environment. At $t = 70s$ the 50 UAS have successfully formed a 3D connected network.

For comparison purposes, a similar experiment was performed, but using the flocking algorithm in [1] instead of the BELBIC-inspired flocking.

3.4.4 Flocking in 3-dimensional Environment with Obstacles

For this scenario, a set of 3D obstacles were created, with the characteristics contained in the following matrix

$$OBS = \begin{bmatrix} 100 & 110 & 110 & 130 & 130 & 160 \\ 4 & 10 & 0 & 2 & 10 & 5 \\ 10 & 3 & 5 & 4 & 6 & 8 \\ 2 & 3 & 2 & 1 & 3 & 5 \end{bmatrix}$$

The first three rows represent the (X, Y, Z) coordinates, while the last row represents the obstacle's radius.

Figure 3.20 shows the 50 UAS at $t = 0s$, in their initial random positions, moving towards the area where the obstacles are located. The UAS form the 3D connected network after completely passing all the obstacles, as can be seen in Figure 3.21.

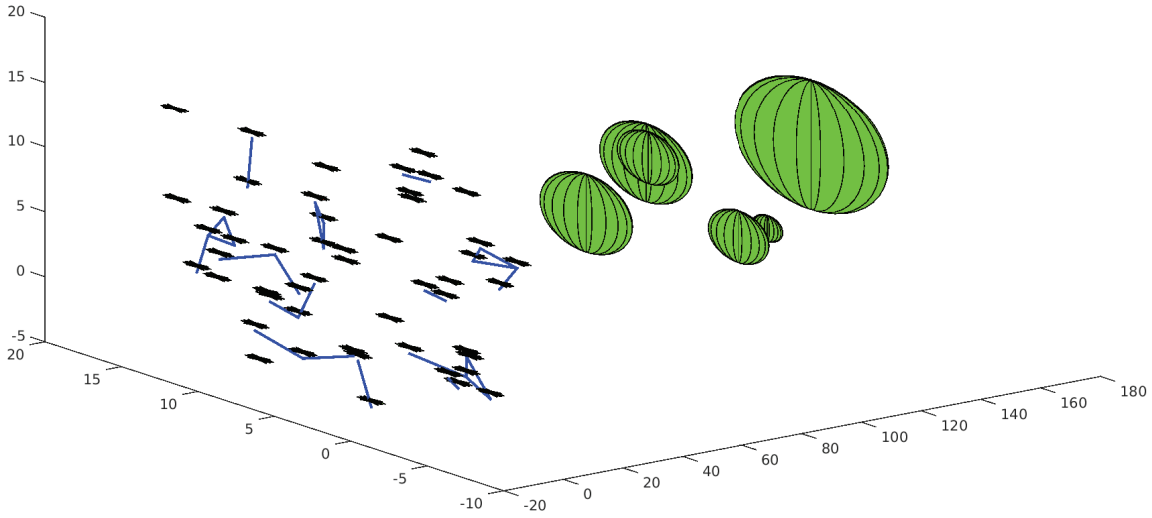

Figure 3.20: Simulation of the BELBIC-inspired flocking algorithm for MAS evolving in an environment with obstacles. At $t = 0s$, the 50 UAS are randomly distributed. The obstacles appear as spheres of different sizes.

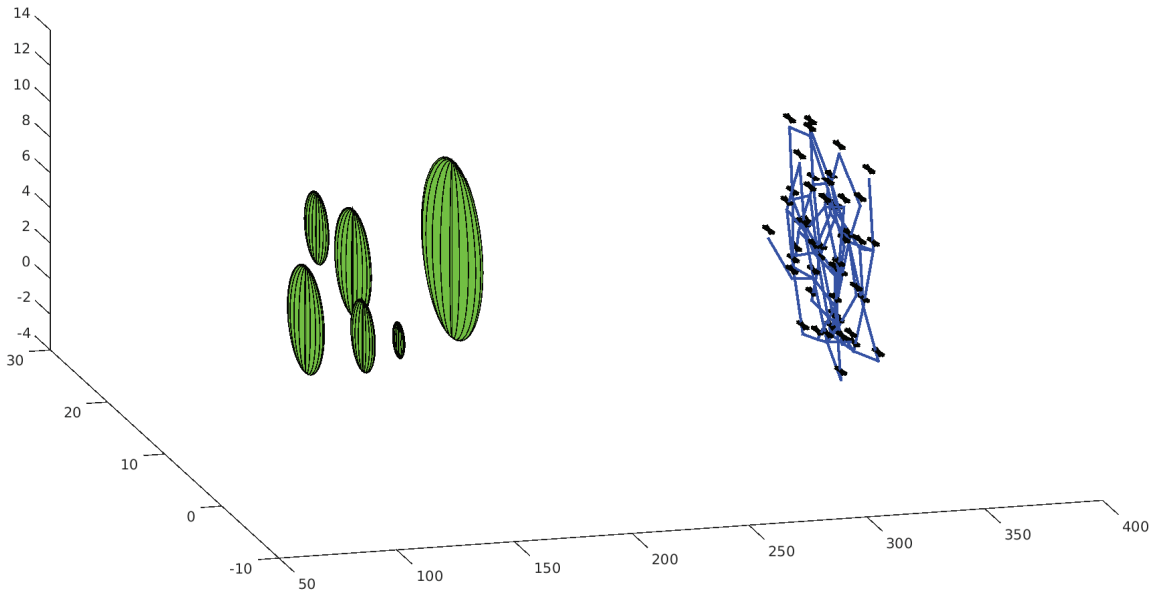

Figure 3.21: Simulation of the BELBIC-inspired flocking algorithm for MAS evolving in an environment with obstacles. At $t = 70s$, the 50 UAS have successfully navigated through the obstacles, without any collision, and have formed a 3D connected network.

For comparison purposes, a similar experiment was performed, but using the flocking algorithm in [1] instead of the BELBIC-inspired flocking.

3.5 Experimental Results

This section presents experimental results showing the performance of the BELBIC-inspired flocking for MAS. Figure 3.22 shows the data flow of the proposed algorithm implementation, which considers the stabilization of four quad rotorcraft UAS.

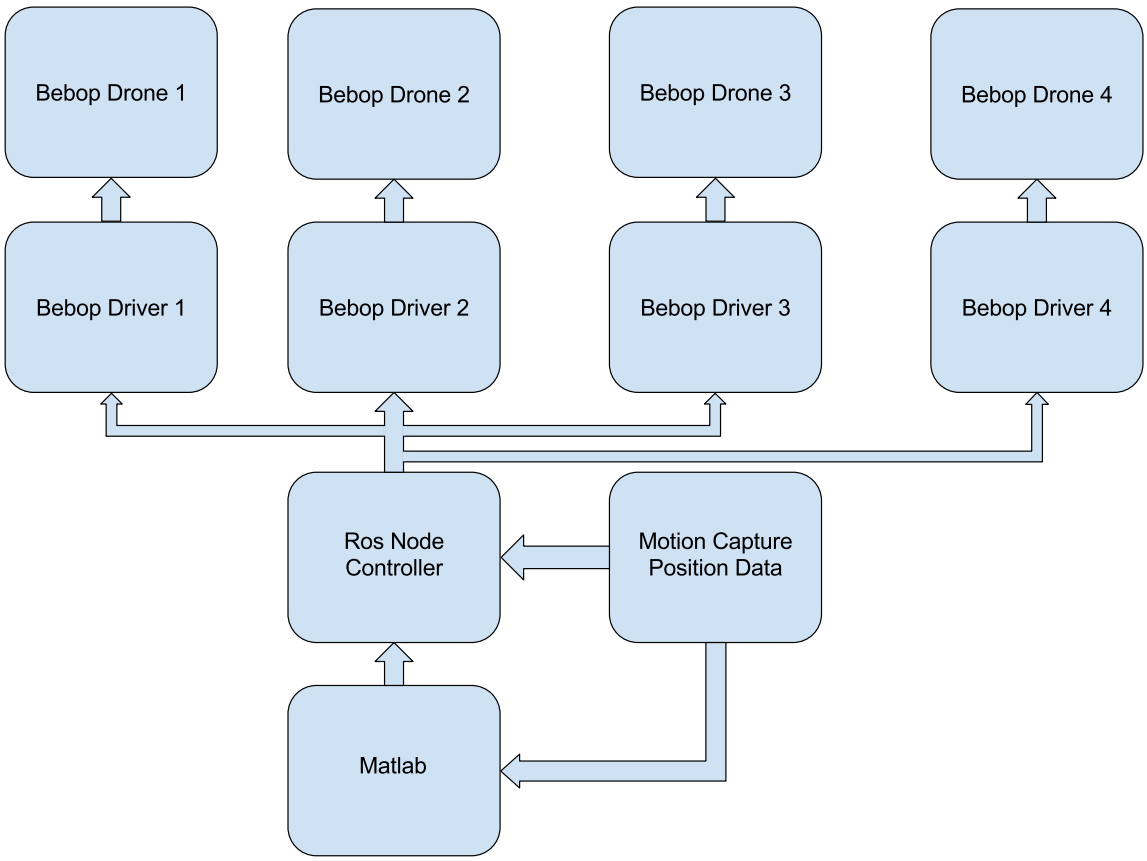

Figure 3.22: Data flow of the proposed algorithm implementation.

3.5.1 Experimental Testbed

The platform implemented for validation of the proposed algorithm is available at the Unmanned Systems Laboratory (USL) from the University of Nevada - Reno [83], see Figure 3.23. The Base Station of this testbed runs Ubuntu 14.04 OS and the Robot Operating System (ROS) environment in combination with Matlab. The proposed algorithms were coded in C/C++ and Matlab, and were executed at 100Hz. The aerial robots correspond to the Bebop drone, manufactured by Parrot.

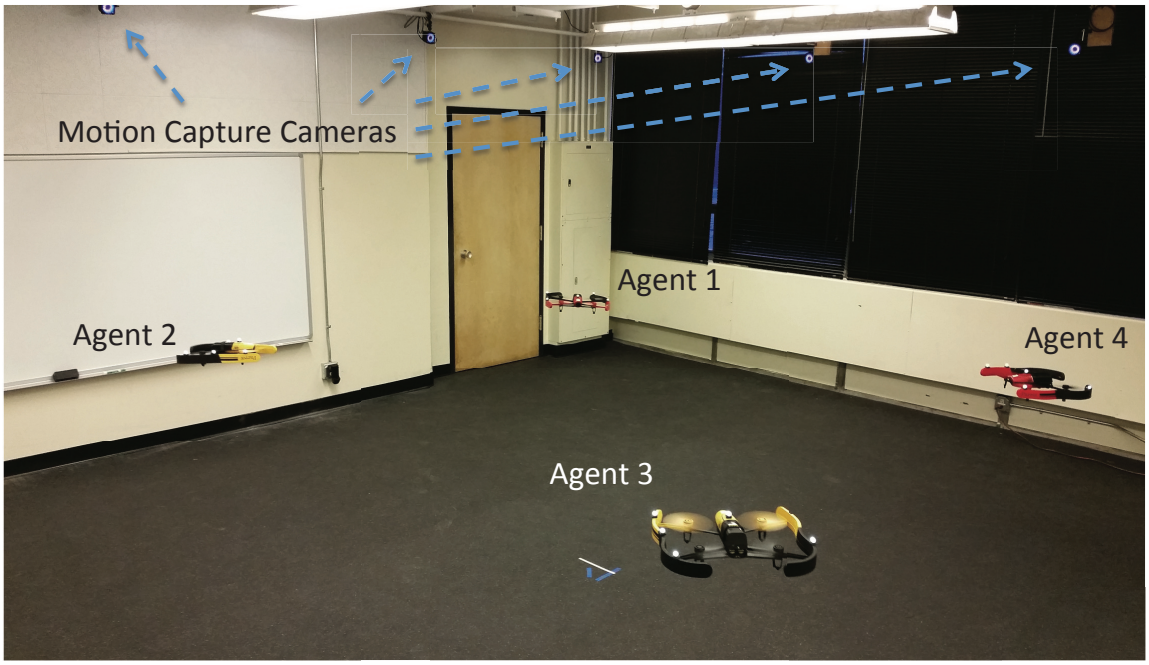

Figure 3.23: Experimental testbed: a set of four drones, a motion capture system, Ground Station computers, and WiFi links.

The 3-dimensional position of the UAS is obtained by means of a Motion Capture System (MCS) manufactured by OptiTrack. The information provided by the MCS is reported to the OptiTrack Interface PC by means of a Gigabyte Ethernet connection. Next, this information is sent to the Base Station PC by means of an Ethernet connection. The Base Station computer uses this information to execute

the BELBIC algorithm and to calculate the control signals, which are then sent to the Bebop platforms by means of a WiFi link.

3.5.2 Real-time Experiments

The ultimate goal of this experimental application is to maintain a satisfactory flocking of the MAS, even when the model of the UAS is uncertain, and when unknown external factors affect the performance of the agents. In this experiment, a total of four UAS were employed, with initial velocities equal to zero, and positions randomly distributed in a squared area. The following parameters are used for the experiment: $r = 1.2d_\alpha$, $d^{'} = 0.6d_\alpha$, $r^{'} = 1.2d^{'}$, while $d_\alpha = 2m$. For the σ-norm the parameter $\epsilon = 0.1$, for $\phi(z)$ the parameters $a = b = 5$, for the bump function $\phi_\alpha(z)$ we used $h = 0.2$, and for $\phi_\beta(z)$ we used $h = 0.9$.

Figure 3.24 plots the positions of the four UAS in (X, Y, Z) axis. It can be seen that all the agents successfully accomplishing the objective of our experimental application.

Figure 3.25 plots the velocities of the four UAS in (X, Y, Z) axis. It can be seen that all the agents agreed on the same speed, successfully accomplishing the objective of our experimental application.

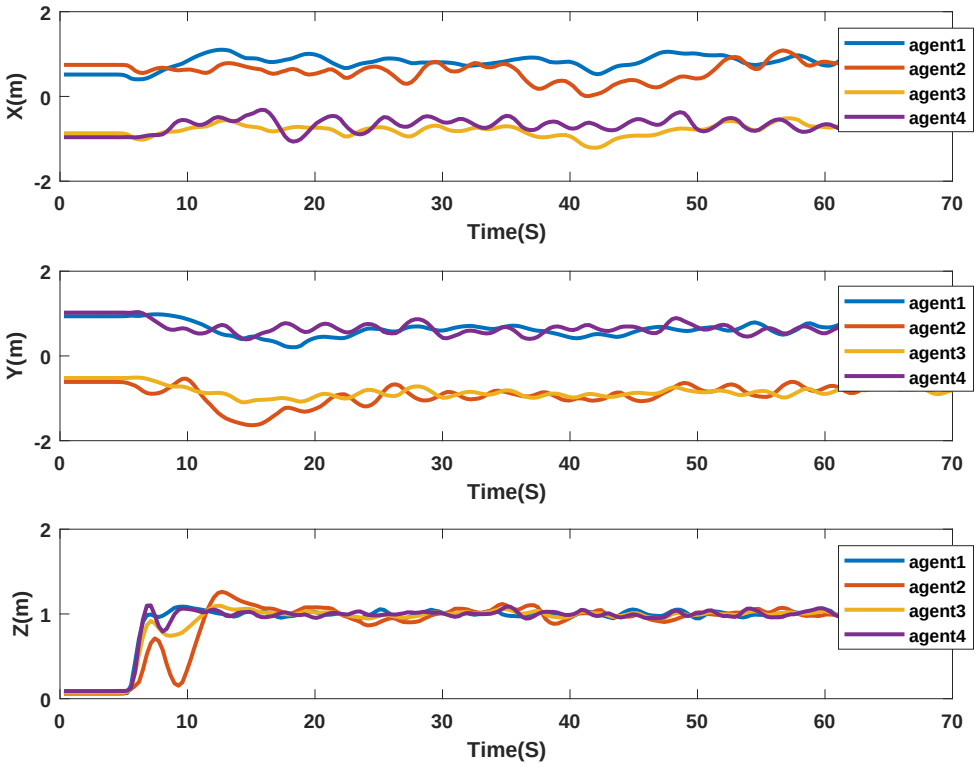

Figure 3.24: BELBIC-inspired flocking of MAS: agreement in UAS positions in (X, Y, Z) axis.

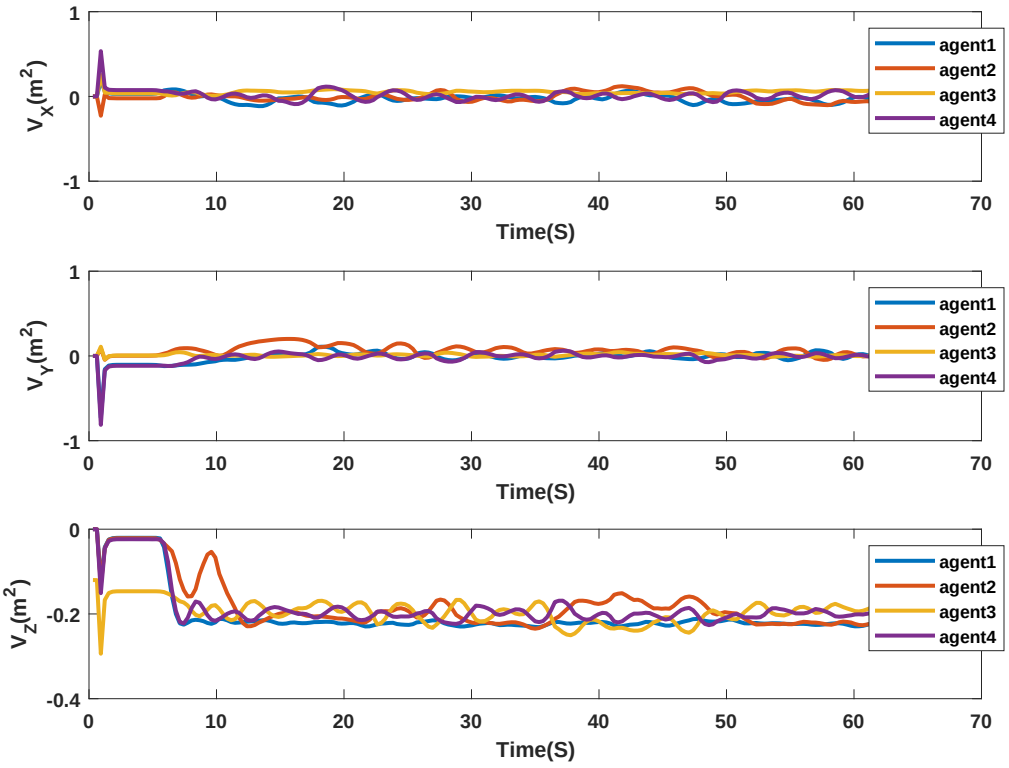

Figure 3.25: BELBIC-inspired flocking of MAS: consensus in UAS velocities in (X, Y, Z) axis.

3.6 Conclusions

A biologically-inspired intelligent controller based on a novel architecture of emotional learning in mammal's brain was proposed for flocking control of MAS. The methodology, which is given the name BELBIC-based flocking, was designed and implemented in the real-time coordination of multiple UAS platforms operating in presence of system uncertainties and dynamic environment. In addition, the convergence analysis of the proposed approach has been studied. Numerical and experimental results of the BELBIC-inspired flocking demonstrated the effectiveness of the proposed approach, as well as its applicability to real-time systems.

3.7 Appendix A

3.7.1 Non-adapting Phase

Our goal is to investigate the output of the system in non-adapting phase (i.e., when the system completes its learning process) so the equations (3.14) and (3.15) which are the updating rules of Amygdala and Orbitofrontal Cortex, respectively, should be taken into consideration. In addition we make an assumption that the max function in equation (3.14) could be neglected. By substituting (3.12) and (3.13) in equation (3.11) the output of the model could be defined as follows:

$$\begin{aligned} MO &= \sum_{l} V_l SI_l - \sum_{l} W_l SI_l \\ &= \sum_{l} \left(V_l - W_l\right) SI_l \end{aligned} \tag{3.23}$$

$$\begin{aligned}\Delta V_l =& K_v SI_l \left(ES - \sum_l A_l \right) \\ =& K_v SI_l \left(ES - \sum_l V_l SI_l \right)\end{aligned} \tag{3.24}$$

$$\begin{aligned}\Delta W_l =& K_w SI_l \left(MO - ES \right) \\ =& K_w SI_l \left(\sum_l \left(V_l - W_l \right) SI_l - ES \right)\end{aligned} \tag{3.25}$$

When the learning process is completed (i.e., after system completes its learning process) the variations of the weights of Amygdala (ΔV_l) and Orbitofrontal Cortex (ΔW_l) will be equal to zero (i.e., $\Delta V_l = \Delta W_l = 0$). With the assumption of $SI_l \neq 0$ the following holds:

$$\begin{aligned}K_v SI_l \left(ES - \sum_l V_l SI_l \right) =& 0 \\ \Rightarrow \sum_l V_l SI_l =& ES\end{aligned} \tag{3.26}$$

$$\begin{aligned}K_w SI_l \left(\sum_l \left(V_l - W_l \right) SI_l - ES \right) =& 0 \\ \Rightarrow \sum_l \left(V_l - W_l \right) SI_l =& ES \\ \Rightarrow \sum_l W_l SI_l =& 0\end{aligned} \tag{3.27}$$

By substituting (3.26) and (3.27) in equation (3.23) the model output in non-adapting phase will be as follows:

$$MO = \sum_l \left(V_l - W_l \right) SI_l = \sum_l V_l SI_l = ES \tag{3.28}$$

3.7.2 Main Proof

Considering the results obtained in Subsection 3.7.1 the following should be achieved:

$$MO_i \rightarrow ES_i \tag{3.29}$$

Let's considering the V_i^* is the weight of Amygdala for each agent i when the system has been learned and let's $\hat{ES_i}$ be the Emotional Signal for each agent i during the adaptation phase. The following hold:

$$ES_i = V_i^* SI_i \;\; and \;\; \hat{ES_i} = V_i SI_i \tag{3.30}$$

$$\Delta V_i(k) = K_v SI_i \max\left(0, ES_i - \hat{ES_i}\right) \tag{3.31}$$

We will investigate the results of the following two cases:

I. $ES_i - \hat{ES_i} \geq 0$

II. $ES_i - \hat{ES_i} < 0$

Considering the case I., the proof can be achieved as follows:

$$\begin{aligned}
\Delta V_i(k) =& K_v SI_i \max\left(0, ES_i - \hat{ES_i}\right) \\
=& K_v SI_i \left(ES_i - \hat{ES_i}\right) \\
=& K_v SI_i \left(V_i^* SI_i - V_i SI_i\right) \\
=& K_v SI_i \left(V_i^* - V_i\right) SI_i \\
=& K_v SI_i \tilde{V_i} SI_i \\
=& K_v \tilde{V_i} \left(SI_i\right)^2
\end{aligned} \tag{3.32}$$

where $\tilde{V}_i = V_i^* - V_i$.

$$\begin{aligned} V_i(k+1) =& V_i(k) + \Delta V_i(k) \\ \tilde{V}_i(k+1) =& V^* - V_i(k) - \Delta V_i(k) \\ =& \tilde{V}_i(k) - K_v \tilde{V}_i (SI_i)^2 \\ =& \left[1 - K_v (SI_i)^2\right] \tilde{V}_i(k) \end{aligned} \tag{3.33}$$

Considering the case II., it is obvious that when $ES_i - \hat{ES}_i < 0$ the max function in equation (3.31) will force the adaptation in Amygdala to stop and the following hold:

$$\begin{aligned} \Delta V_i(k) =& 0 \\ V_i(k+1) =& V_i(k) \\ \tilde{V}_i(k+1) =& \tilde{V}_i(k) \end{aligned} \tag{3.34}$$

The proof can be achieved as follows:

$$\begin{aligned} \Delta W_i(k) =& K_w SI_i (MO_i - ES_i) \\ =& K_w SI_i (V_i SI_i - W_i SI_i - V_i^* SI_i) \\ =& K_w SI_i (-(V_i^* - V_i) SI_i - W_i SI_i) \\ =& K_w SI_i \left(\left(-\tilde{V}_i - W_i\right) SI_i\right) \\ =& - K_w \tilde{W}_i (SI_i)^2 \end{aligned} \tag{3.35}$$

where $\tilde{V}_i = V_i^* - V_i$ and $\tilde{W}_i = \tilde{V}_i + W_i$.

$$\begin{aligned} W_i(k+1) =& W_i(k) + \Delta W_i(k) \\ \tilde{W}_i(k+1) =& \tilde{V}_i(k+1) + W_i(k+1) \\ =& \tilde{V}_i(k) + W_i(k) + \Delta W_i(k) \\ =& \tilde{W}_i(k) - K_w \tilde{W}_i (SI_i)^2 \\ =& \left[1 - K_w (SI_i)^2\right] \tilde{W}_i(k) \end{aligned} \tag{3.36}$$

Chapter 4

A Biologically-Inspired Distributed Fault Tolerant Flocking Control for Multi-Agent System in Presence of Uncertain Dynamics and Unknown Disturbance

In this paper, the problem of flocking for Multi-Agent Systems (MAS) in presence of system uncertainties and unknown disturbances is investigated. A biologically-inspired novel distributed resilient controller based on a computational model of emotional learning in the mammalian brain is proposed. The methodology, known as Brain Emotional Learning Based Intelligent Controller (BELBIC), integrates a

resilience mechanism with multi-objective properties into the distributed flocking control strategy. The developed strategy adopts the learning capabilities of BELBIC with the flocking that makes it a very promising approach, especially while dealing with system uncertainties and/or unknown disturbances. Furthermore, the low computational complexity of the designed method makes it very suitable for practical implementation in real-time applications. Eventually, the effectiveness of the developed intelligent resilient distributed flocking control approach is demonstrated through the several simulation scenarios.

4.1 Introduction

Coordination control of Multi-Agent Systems (MAS) has received a great deal of interests from diverse research communities during the past decade, see [1,3,6,11] and the references therein. Meanwhile, flocking design for MAS has also attracted enormous interests [88]. Flocking is a fact in which a large number of self-propelled entities form the coordinated and collective motion. This phenomenon is exhibited by many living beings such as birds, fish, bacteria, and insects [15].

In the pioneering work [7], three fundamental rules for simulating the flocking behavior are introduced, specifically, *separation*, *alignment*, and *cohesion*. Recently, diverse researchers spent tremendous efforts to improve the flocking behavior of MAS [6,17,88].

However, most of those works only considered the agents with accurate model. Inherently, MAS contains several critical problems such as system uncertainties and unknown disturbances from harsh environment. The performance of current flocking algorithms could be significantly degraded even fail to maintain the sys-

tem stability due to the dispersion of disturbances and/or uncertainties. Aiming at addressing the issues arising from practical harsh conditions, diverse research efforts have been proposed. For example, the authors in [27] proposed a distributed target tracking and estimation algorithm using multiple autonomous agents with uncertain dynamics. Robust consensus tracking for MAS with disturbances and unmodeled dynamics has been investigated in [28]. The neural adaptive leader-follower flocking control of MAS has been addressed in [30]. The authors in [16], investigated the flocking of MAS with communication delays in a noisy environment. Adaptive flocking of nonlinear MAS with uncertain parameters is addressed in [19]. Closely related, the leader-follower flocking problem with a moving leader and parametric uncertainties is studied in [18]. In [20], flocking with connectivity preservation of MAS subject to external disturbances has been developed. Most recently, adaptive architectures for resilient control of networked MAS in the presence of misbehaving agents has been proposed in [89]. Also, the authors in [90] presented a resilient formation control for mobile robot teams in a non-cooperative manner.

However, most of these recent approaches highly depend on entire or partial information about the system dynamics. Also, the developed learning-based approaches are computationally complex and not appropriate for real-time implementation. Moreover, the practical impacts from unknown disturbances have not been considered in most of the existing flocking algorithms [1, 17, 18]. To overcome these deficiencies, a novel low-computational biologically-inspired resilient distributed flocking control algorithm will be proposed in this paper. This proposed method cannot only guarantee the desired flocking performance, and also handle the system uncertainties and unknown disturbances through a novel low-computational biologically-inspired learning approach.

In recent years, intelligent approaches have been extensively utilized for successfully solving diverse complex problems [23,91–95]. Among them, biologically-inspired intelligent approaches have been received tremendous interests by many researchers because of their inherent potential to deal with computationally complex systems. One of such approaches which adopts the computational model of emotional learning introduced in [40] is Brain Emotional Learning Based Intelligent Controller (BELBIC). This methods mimics parts of the mammalian brain which are responsible of producing the emotion i.e., the *amygdala*, *orbitofrontal cortex*, *thalamus*, and *sensory input cortex*. BELBIC based methodologies have been demonstrated to be a reliable solution while dealing with uncertainties [43], [41]. *Sensory Inputs* (SI) and *Emotional Signal* (ES) are the two main inputs to the BELBIC. This biologically-inspired intelligent controller is an appealing strategy for solving the multi-objective problems in real-time applications due to the flexibility in defining both SI and ES [42]. Moreover, BELBIC has a low computational complexity on the order of $O(n)$.

The main contribution of this paper is to develop Resilient BELBIC (R-BELBIC) as an intelligent distributed resilient controller for flocking of MAS in presence of system uncertainties and unknown disturbances. To the best of the authors' knowledge, this is the first time that R-BELBIC is developed and implemented for accomplishing MAS flocking. In this paper, we focus on the problem of intelligent distributed resilient flocking of MAS. Our proposed methodology builds on the distributed flocking algorithm. In order to guarantee resilience, we employed the computational model of emotional learning in the mammalian limbic system, i.e., BELBIC with context. We then demonstrated that the developed control law achieves resilient flocking of MAS through both computer-aid simulation and experimental tests. Ultimately, individual robots follow the developed con-

trol strategy that ensures resilience, and use the R-BELBIC flocking algorithm to guarantee the suitable behavior even in the presence of uncertainties and disturbances due to the learning capability of the proposed approach. Furthermore, R-BELBIC has a very low computational cost which makes it a promising method in a real-time application. The developed R-BELBIC design provides an intelligent distributed resilient flocking controller with multi-objective properties i.e., control effort optimization, uncertainty handling, and noise/disturbance rejection, while maintaining the system complexity in a practically achievable limit. In order to demonstrate the effectiveness of the proposed approach, a set of numerical simulations are provided. Additionally, a comparison between the developed approach and conventional flocking strategies is provided, where it is possible to observe the resilient flocking strategy attained with R-BELBIC.

The rest of the paper is organized as follows. Section 4.2 presents the problem formulation and preliminaries about MAS flocking and BELBIC. Our main contribution is given in Section 4.3, which consists of an intelligent distributed resilient flocking control strategy of MAS based on R-BELBIC. Section 4.4 presents numerical simulation results and Section 4.5 presents an experimental validation of developed method. The conclusion of the paper and future directions of our work are provided in Section 4.6.

4.2 Problem Formulation and Preliminaries

In this section, some preliminaries are provided and the problem formulation is briefly discussed. First, the dynamic of the MAS is given, then, flock topology is modeled by means of a dynamic graph. Next, BELBIC model is introduced, and

ultimately, the problem is formulated.

4.2.1 MAS Flocking

Consider n agents, with double integrator dynamics, evolving in an m dimensional space $(m = 2, 3)$. The equations of motion of each agent can be described as

$$\begin{cases} \dot{q}_i = p_i \\ \dot{p}_i = u_i \qquad i = 1, 2, ..., n \end{cases} \tag{4.1}$$

where u_i is the control input, q_i is the position, and p_i is the velocity of the agent i, respectively. The flock topology is modelled by means of a dynamic graph $\mathcal{G}(\upsilon, \varepsilon)$ which consists of a set of vertices $\upsilon = \{1, 2, ..., n\}$ representing the agents, and edges $\varepsilon \subseteq \{(i, j) : i, j \in \upsilon, j \neq i\}$ representing the communication link between a pair of agents. The neighborhood set of agent i is described by

$$N_i^{\alpha} = \{j \in \upsilon_{\alpha} :\| q_j - q_i\| < r, j \neq i\} \tag{4.2}$$

where r is a positive constant expressing the range of interaction between agents i and j, and $\| \cdot \|$ is the Euclidean norm in $\mathbb{R}^{\mathrm{m}}$. The geometric model of the flock, i.e., the α-lattice [1] is accomplished by solving the following set of algebraic constraints

$$\| q_j - q_i\| = d \quad \forall j \in N_i^{\alpha} \tag{4.3}$$

where the positive constant d describes the distance between two neighbors i and j. In order to resolve the singularity problem caused at $q_i = q_j$ in the collective potential function, equation (4.3) can be rewritten as

$$\| q_j - q_i\|_{\sigma} = d_{\alpha} \quad \forall j \in N_i^{\alpha} \tag{4.4}$$

where $d_\alpha = \| d\|_\sigma$, and $\| \cdot\|_\sigma$ is the σ-norm which is represented by $\| z\|_\sigma = \frac{1}{\epsilon}[\sqrt{1+\epsilon\| z\|^2} - 1]$, and $\epsilon > 0$ is a positive constant. For a vector z, the σ-norm represents a map from $\mathbb{R}^{\mathrm{m}}$ to $\mathbb{R} \geq 0$. The new map $\| z\|_\sigma$ is differentiable everywhere, while the Euclidean norm $\| z\|$ is not differentiable at $z = 0$.

Considering the above conditions, a smooth collective potential function is defined as

$$V(q) = \frac{1}{2}\sum_i \sum_{j \neq i} \psi_\alpha(\| q_j - q_i\|_\sigma)$$

where $\psi_\alpha(z)$ is a smooth pairwise potential function determined by $\psi_\alpha(z) = \int_{d_\alpha}^{z} \phi_\alpha(s)ds$, with $\phi_\alpha(z) = \rho_h(z/r_\alpha)\phi(z - d_\alpha)$, and $\phi(z) = \frac{1}{2}[(a+b)\sigma_1(z+c) + (a-b)]$, with $\sigma_1(z) = z/\sqrt{1+z^2}$. Notice that $\phi(z)$ is an uneven sigmoidal function with $0 < a \leq b$, $c = |a-b|/\sqrt{4ab}$ to guarantee that $\phi(0) = 0$. In addition, $\rho_h(z)$ is a scalar bump function which smoothly varies between $[0, 1]$. A possible choice for determining $\rho_h(z)$ is [1]

$$\begin{cases} 1, & z \in [0, h) \\ \frac{1}{2}\left[1 + \cos\left(\pi \frac{(z-h)}{(1-h)}\right)\right], & z \in [h, 1] \\ 0, & \text{otherwise} \end{cases} \tag{4.5}$$

The flocking control algorithm $u_i = u_i^\alpha + u_i^\gamma$ introduced in [1] makes all agents to form an α-lattice configuration, while at the same time avoiding any collisions, and following a specific trajectory. The algorithm consists of the following main components: (i) u_i^α which is the interaction term between two α-agents i and j, (ii) u_i^γ which is a goal term and represents a distributed navigational feedback component. Each one of these components is explicitly expressed as follows

$$u_i^\alpha = c_1^\alpha \sum_{j \in N_i^\alpha} \phi_\alpha(\| q_j - q_i\|_\sigma)\mathbf{n}_{i,j} + c_2^\alpha \sum_{j \in N_i^\alpha} a_{ij}(q)(p_j - p_i)$$

$$u_i^{\gamma} = -c_1^{\gamma}\sigma_1(q_i - q_r) - c_2^{\gamma}(p_i - p_r)$$

where c_1^{α}, c_2^{α}, c_1^{γ}, and c_2^{γ} are positive gains. The pair (q_r, p_r) represents the γ-agent, which is the virtual leader of the flock and is described as

$$\begin{cases} \dot{q}_r = p_r \\ \dot{p}_r = f_r(q_r, p_r) \end{cases} \tag{4.6}$$

The term $\mathbf{n}_{i,j}$ is a vector determined by

$$\mathbf{n}_{i,j} = \frac{q_j - q_i}{\sqrt{1 + \epsilon \| \, q_j - q_i \|^2}}$$

The term $a_{ij}(q)$ is the element of the spatial adjacency matrix $A(q)$, which is expressed as $a_{ij}(q) = \rho_h(\| \, q_j - q_i \|_{\sigma} / r_{\alpha}) \in [0, 1], j \neq i$. In this equation $r_{\alpha} = \| \, r \|_{\sigma}$, $a_{ii}(q) = 0$ for all i and q.

4.2.2 Brain Emotional Learning-Based Intelligent Controller

BELBIC was first introduced in [41], and since then it has been utilized in different applications exhibiting very satisfactory results. The authors in [96] applied this controller for performing the automatic speed control of an asymmetrical six-phase induction motor. A BELBIC load-frequency control of interconnected power system has been addressed in [97]. The authors in [43] successfully applied this controller for the attitude control of a quad rotorcraft. More recently, BELBIC was employed for the navigation of an Unmanned Ground Vehicle in [74]. Closely related, the authors in [75] addressed the optimal bi-objective structure emotional learning for a Dynamic Voltage Restorer.

Furthermore, several research groups have presented experimental applications based on the real-time implementation of BELBIC. For example, the authors

in [87] implemented this strategy on a Field-Programmable Gate Array for controlling an overhead traveling crane. A Digital Signal Processor board was employed to practically implement BELBIC for controlling an interior permanent magnet synchronous motor drive in [86]. The authors in [46] experimentally applied this strategy for controlling the speed of a digital servo system.

The learning capability of an intelligent system is a pivotal characteristic which distinguishes it from a conventional system [41]. It is essential for any learning methodology to have an appraisal mechanism for determining the operational condition of the system to cope with the changing environment. One category of appraisal mechanisms is based on the emotional cues [98]. This evaluation mechanism determines how external stimuli could affect the capability of the system to effectively function in short-term and also to retain its long-term prospects for survival. Emotional Learning, is a learning technique that is based on emotional appraisals.

The Sensory Inputs (SI) and the Emotional Signal (ES) are the two inputs to the BELBIC model. The output of this model can be described by the following equation

$$MO = \sum_{l} A_l - \sum_{l} OC_l \tag{4.7}$$

which is determined by subtracting the Orbitofrontal Cortex outputs (OC_l) from the Amygdala outputs (A_l). In this equation, l represents the number of sensory inputs.

The outputs of the Amygdala and Orbitofrontal Cortex are computed by adding all their corresponding nodes, see Figure 4.1, where the outputs of each node are expressed as

$$A_l = V_l SI_l \tag{4.8}$$

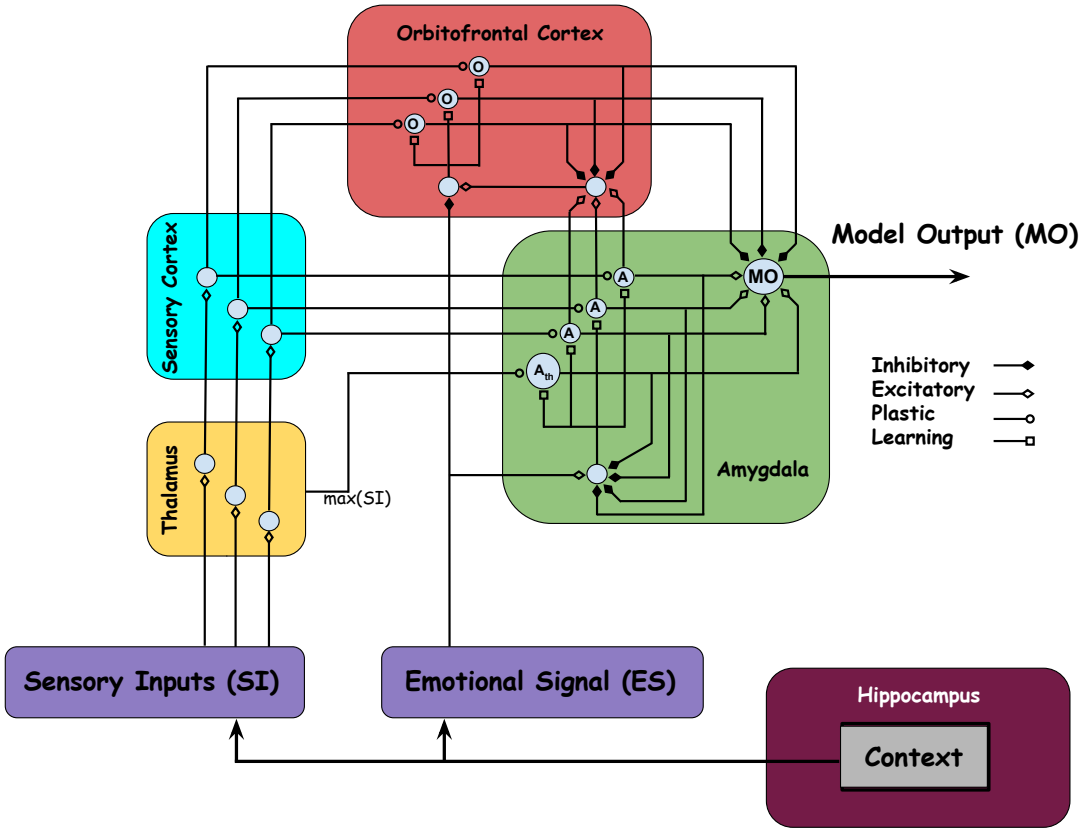

Figure 4.1: Computational model of emotional learning (Consisting of three main part *Amygdala*, *Orbitofrontal Cortex*, and *Hippocampus*).

$$OC_l = W_l SI_l \tag{4.9}$$

where W_l is the weight associated to the Orbitofrontal Cortex, V_l is the weight associated to the Amygdala, and SI_l is the l^{th} sensory input.

In order to update V_l and W_l we use

$$\Delta V_l = K_v SI_l \max\left(0, ES - \sum_l A_l\right) \tag{4.10}$$

$$\Delta W_l = K_w SI_l \left(MO - ES\right) \tag{4.11}$$

where K_v and K_w are the learning rates of the Amygdala and the Orbitofrontal Cortex, respectively.

There is another input to the Amygdala, which directly comes from the Tha-

lamus. This input is computed by finding the maximum of all SI and can be described as

$$A_{th} = V_{th} \max(SI_l) \tag{4.12}$$

Several techniques have been employed for tuning the BELBIC parameters [41, 43, 47, 79–81]. In this paper, a heuristic approach is employed for determining the BELBIC parameters.

4.2.3 Flocking Control Objectives

The problematic and goal of this research can be summarized as follows. Considering the model uncertainty and disturbances, the new dynamics could be described as follows:

$$\begin{cases} \dot{q}_i = p_i \\ \dot{p}_i = u_i + \chi_i p_i + \Gamma_i \quad , i = 1, 2, ..., n \end{cases} \tag{4.13}$$

where $\{u_i, q_i, p_i\} \in \mathbb{R}^{\mathrm{m}}$ are same as (equation 4.1) and $\chi_i \in \mathbb{R}^{\mathrm{m}\times\mathrm{m}}$ is the unknown uncertain parameters matrix and $\Gamma_i \in \mathbb{R}^{\mathrm{m}}$ is an unknown vector of disturbances respectively.

The objective is then, to design an intelligent distributed resilient controller, in such a way that the motion of the agents in the flock exhibits an emergent behavior caused by a set of simple rules, which are executed by each agent independently, and do not require any central coordination. Moreover, the proposed controller should keep a low level of complexity, in order to be practically implementable in the real-time flocking of MAS.

Ultimately, the proposed methodology will enable the flock to track a virtual leader (i.e., the γ-agent), while avoiding collisions with every other agent in the

group. Furthermore, the proposed R-BELBIC flocking strategy should be able to satisfy multiple control objectives, such as (i) minimization of the control effort and energy expenses, (ii) resilient against environmental noise/disturbances, and (iii) model uncertainty handling. All of these previous requirements should be attained without increasing the complexity of the system.

Remark 5. *In practical MAS flocking control systems, due to the dynamic environment, autonomous agents are commonly suffering from different type of faults. Since the dynamics of the quad rotorcrafts employed for real-time experimental evaluation are unknown, and since we did our experiment in a small laboratory where we have uncertain disturbances affecting the MAS due to the close operation of neighboring quad rotorcrafts (i.e., wind disturbances). Therefore, in this paper, our analysis is focused on developing an intelligent distributed resilient controller for flocking of MAS in presence of system uncertainties and unknown disturbances from complex environment. One can extend our results also to handle other resources of faults such as network imperfections [99], actuator faults [100], etc.*

Remark 6. *Multi-agent systems should operate more economical, more reliable, and much safer, to be practically applicable in real world scenarios such as tracking, search and rescue, underwater explorations, etc [101–104]. Therefore, improving the performance of MAS by leveraging the information about the environment (i.e., context) needs to be studied. It is worth mentioning that, to the best of authors' knowledge, this is the first attempt to integrate the Hippocampus to provide the context to the BELBIC model to develop Resilient BELBIC (RBELBIC) as an intelligent distributed resilient controller for flocking of MAS in presence of system uncertainties and unknown disturbances. Also, R-BELBIC is implemented in real-time for accomplishing MAS flocking by employing multiple UAS. In addition, our model also is an appealing strategy for solving the multi-objective problems in real-time applications due to the flexibility in defining both SI and ES [41]. Moreover,*

R-BELBIC has a low computational complexity. Furthermore, we provided the Lyapunov stability analysis to demonstrate that our proposed methodology guarantees the convergence of the designed control signals as well as maintain the system stability during the learning. Hence, it is of both practical and theoretical significance to study different MAS problems by employing R-BELBIC.

4.3 Intelligent Distributed Resilient Flocking Control of MAS Using R-BELBIC

One component has been missing from the model described in subsection 4.2.2. Hippocampus is an important component in emotional conditioning, and it is responsible for supplying the system with a context [98]. Based on different theories and models, many roles and functionalities have been assigned to the hippocampus. The hippocampus has been known by many functionalities, including but not limited to: (i) spatial navigation, (ii) establishing the long-term memory and (iii) the contextual representations [98]. Also, mapping of the environment based on the environmental cues is one of the hippocampus responsibilities [98]. Furthermore, the comparison between stored regularities and actual stimuli is one of the important functionalities of the hippocampal system [98].

Note, our focus is on hippocampus effects on the amygdala model. Here, we investigated the hippocampus effects on the system by adding the *Context* block to the system. "*Context is a feature that is frequently defined in the negative, as any stimuli that is not directly involved in the present experiment, stimuli that somehow encode the entire situation, rather than individual features, or as stimuli that are not being manipulated by the experimenter*" [98].

Figure 4.2 illustrate the R-BELBIC architecture implemented in this work.

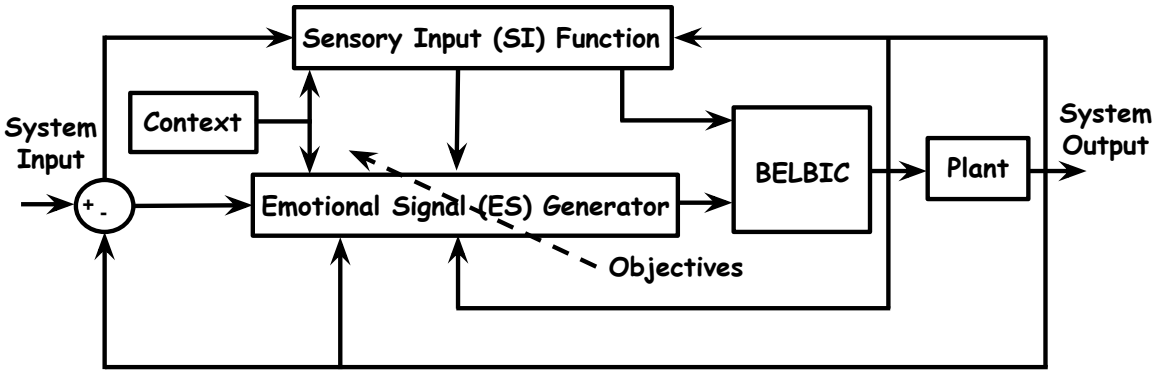

Figure 4.2: R-BELBIC in the control loop.

The general form of the SI and the ES signals are as follow:

$$SI = G\left(t, y, e, u, r, CON\right) \tag{4.14}$$

$$ES = F\left(t, y, e, u, r, CON\right) \tag{4.15}$$

where t, e, u, y, and r, are time, system error, control effort, system output, and system input, respectively. The term CON is a new term which represents the Context. Specific control objectives can be implicitly determined by choosing the appropriate ES. For example, it is possible to choose the ES in such a way to optimize the energy expense, to preserve network connectivity, or to attain superior target tracking performance, among others.

Remark 7. *Essentially, R-BELBIC is producing its actions based on the Emotional signal (ES) and Sensory input (SI). The current situation of the system is represented by the received sensory inputs and the emotional signal represents the condition of the system by considering the particular objective of interest. Therefore, by assigning different values to SI and ES, they would have different impacts on the performance of the whole system. In other words, these functions provide the design freedom allowing different priorities for*

flocking objective such as tracking, varying objectives, etc. However, our objective in this work is to handle system uncertainties and unknown disturbances, while also considering the perfect flocking behavior.

The R-BELBIC-inspired flocking focuses on handling the system uncertainties and disturbances (resilient), while at the same time minimizing the control effort and the energy of the flock. By doing this, we will accomplish a practical controller well suited for implementation in real-time applications. In order to fulfill the above-mentioned objectives, the ES is designed in such a way that the increase in control effort will express a *negative emotion*, e.g., stress. Therefore, the more stress emanated from the system, the more the system performance will be considered as not satisfactory. In addition, by adding the context term, it becomes resilient dealing with system uncertainties and disturbances.

The utilized sensory inputs (SI_i) and emotional signal (ES_i), for each agent i, are expressed as follows

$$SI_i = K^{\alpha}_{SI,i} u^{\alpha}_i + K^{\gamma}_{SI,i} u^{\gamma}_i + K^{CON}_{SI,i} \Upsilon_i \left(u_i, q_i, p_i\right) \tag{4.16}$$

$$ES_i = K^{\alpha}_{ES,i} u^{\alpha}_i + K^{\gamma}_{ES,i} u^{\gamma}_i + K^{CON}_{ES,i} \Upsilon_i \left(u_i, q_i, p_i\right) \tag{4.17}$$

where $K^{\alpha}_{SI,i}$, $K^{\gamma}_{SI,i}$, $K^{CON}_{SI,i}$, $K^{\alpha}_{ES,i}$, $K^{\gamma}_{ES,i}$, $K^{CON}_{ES,i}$ are positive constants and $\Upsilon_i \left(u_i, q_i, p_i\right)$ is a function representing Context. By assigning different values to these gains, the influence of ES on the system behavior will change. Since this work emphasizes the resilient control of the system, the gains for the context are greater than the others.

In a real-time scenario, the MAS could face unexpected events. The designer can take advantage of the adaptability of the equation (4.17) in order to successfully

achieve different flocking objectives. Then, from the equations (4.16)-(4.17), the BELBIC-inspired flocking of MAS is expressed as

$$\begin{aligned} u_i^{R-BEL} &= \sum_i V_i.SI_i - \sum_i W_i.SI_i \\ &= \sum_i V_i.\left(K_{SI,i}^{\alpha}u_i^{\alpha} + K_{SI,i}^{\gamma}u_i^{\gamma} + K_{SI,i}^{CON}\Upsilon_i\left(u_i, q_i, p_i\right)\right) \\ &\quad - \sum_i W_i.\left(K_{SI,i}^{\alpha}u_i^{\alpha} + K_{SI,i}^{\gamma}u_i^{\gamma} + K_{SI,i}^{CON}\Upsilon_i\left(u_i, q_i, p_i\right)\right) \end{aligned} \tag{4.18}$$

Here $i = 1, ..., n$ and n is the number of agents. By considering the results obtained from Theorem 4 and by substituting the Emotional Signal with equation (4.17) the R-BELBIC flocking model output for MAS could be obtained as follows:

$$MO_i = ES_i = K_{ES,i}^{\alpha}u_i^{\alpha} + K_{ES,i}^{\gamma}u_i^{\gamma} + K_{ES,i}^{CON}\Upsilon_i\left(u_i, q_i, p_i\right) \tag{4.19}$$

which is exactly demonstrating the flocking behavior of the proposed method.

Stability Analysis

The convergence of the weights of Amygdala (V_i) and Orbitofrontal Cortex (W_i) are presented in Theorem 4.

Theorem 4. *Given the R-BELBIC design as* (4.16)–(4.19)*, there exists the positive R-BELBIC tuning parameter, K_{vi}, K_{wi} satisfying*

I. $\left|\left[1 - K_{vi}\left(SI_i\right)^2\right]\right| < 1$

II. $\left|\left[1 - K_{wi}\left(SI_i\right)^2\right]\right| < 1$

such that the MAS's estimated weights of Amygdala (V_i) and Orbitofrontal Cortex (W_i) converge to desired targets asymptotically.

Proof. **Proof of Theorem 4**

In this subsection, to evaluate the performance of the whole system, the behavior of BELBIC main algorithm (4.2.2) and developed R-BELBIC for MAS flocking is investigated. Since the output of the developed model, when the system completes its learning phase, is important for our main design, the adaptations rules of BELBIC is initially used in (4.3) to show the behavior of the BELBIC model. Then, by utilizing the obtained results in (4.3), the behavior of the developed R-BELBIC methodology is studied in (4.3).

Non-adapting Phase

Our goal is to investigate the output of the system in non-adapting phase (i.e., when the system completes its learning process) so the equations (4.10) and (4.11) which are the updating rules of Amygdala and Orbitofrontal Cortex, respectively, should be taken into consideration. In addition we make an assumption that the max function in equation (4.10) could be neglected. By substituting (4.8) and (4.9) in equation (4.7) the output of the model could be defined as follows:

$$\begin{aligned} MO &= \sum_l V_l SI_l - \sum_l W_l SI_l \\ &= \sum_l \left(V_l - W_l\right) SI_l \end{aligned} \tag{4.20}$$

$$\begin{aligned}\Delta V_l &= K_v SI_l \left(ES - \sum_l A_l \right) \\ &= K_v SI_l \left(ES - \sum_l V_l SI_l \right)\end{aligned} \tag{4.21}$$

$$\begin{aligned}\Delta W_l &= K_w SI_l \left(MO - ES \right) \\ &= K_w SI_l \left(\sum_l \left(V_l - W_l \right) SI_l - ES \right)\end{aligned} \tag{4.22}$$

When the learning process is completed (i.e., after system completes its learning process) the variations of the weights of Amygdala (ΔV_l) and Orbitofrontal Cortex (ΔW_l) will be equal to zero (i.e., $\Delta V_l = \Delta W_l = 0$). With the assumption of $SI_l \neq 0$ the following holds:

$$\begin{aligned}K_v SI_l \left(ES - \sum_l V_l SI_l \right) &= 0 \\ \Rightarrow \sum_l V_l SI_l &= ES\end{aligned} \tag{4.23}$$

$$\begin{aligned}K_w SI_l \left(\sum_l \left(V_l - W_l \right) SI_l - ES \right) &= 0 \\ \Rightarrow \sum_l \left(V_l - W_l \right) SI_l &= ES \\ \Rightarrow \sum_l W_l SI_l &= 0\end{aligned} \tag{4.24}$$

By substituting (4.23) and (4.24) in equation (4.20) the model output in non-adapting phase will be as follows:

$$MO = \sum_l \left(V_l - W_l \right) SI_l = \sum_l V_l SI_l = ES \tag{4.25}$$

Main Proof

Considering the results obtained in Subsection 4.3 the following should be achieved:

$$MO_i \rightarrow ES_i \tag{4.26}$$

Let's considering the V_i^* is the weight of Amygdala for each agent i when system has been learned and let's $\hat{ES}_i$ be the Emotional Signal for each agent i during the adaptation phase. The following hold:

$$ES_i = V_i^* SI_i \;\; and \;\; \hat{ES}_i = V_i SI_i \tag{4.27}$$

$$\Delta V_i(k) = K_{vi} SI_i \max\left(0, ES_i - \hat{ES}_i\right) \tag{4.28}$$

We will investigate the results for the following two cases:

I. $ES_i - \hat{ES}_i \geq 0$

II. $ES_i - \hat{ES}_i < 0$

Considering the case I., the proof can be achieve as follows:

$$\begin{aligned}
\Delta V_i(k) &= K_{vi} SI_i \max\left(0, ES_i - \hat{ES}_i\right) \\
&= K_{vi} SI_i \left(ES_i - \hat{ES}_i\right) \\
&= K_{vi} SI_i \left(V_i^* SI_i - V_i SI_i\right) \\
&= K_{vi} SI_i \left(V_i^* - V_i\right) SI_i \\
&= K_{vi} SI_i \tilde{V}_i SI_i \\
&= K_{vi} \tilde{V}_i \left(SI_i\right)^2
\end{aligned} \tag{4.29}$$

where $\tilde{V}_i = V_i^* - V_i$.

$$\begin{aligned} V_i(k+1) &= V_i(k) + \Delta V_i(k) \\ \tilde{V}_i(k+1) &= V^* - V_i(k) - \Delta V_i(k) \\ &= \tilde{V}_i(k) - K_{vi}\tilde{V}_i(SI_i)^2 \\ &= \left[1 - K_{vi}(SI_i)^2\right]\tilde{V}_i(k) \end{aligned} \tag{4.30}$$

Considering the case II., it is obvious that when $ES_i - \hat{ES}_i < 0$ the max function in equation (4.28) will force the adaptation in Amygdala to stop and the following hold:

$$\begin{aligned} \Delta V_i(k) &= 0 \\ V_i(k+1) &= V_i(k) \\ \tilde{V}_i(k+1) &= \tilde{V}_i(k) \end{aligned} \tag{4.31}$$

The proof can be achieve as follows:

$$\begin{aligned} \Delta W_i(k) &= K_{wi}SI_i(MO_i - ES_i) \\ &= K_{wi}SI_i(V_iSI_i - W_iSI_i - V_i^*SI_i) \\ &= K_{wi}SI_i(-(V_i^* - V_i)SI_i - W_iSI_i) \\ &= K_{wi}SI_i\left(\left(-\tilde{V}_i - W_i\right)SI_i\right) \\ &= -K_{wi}\tilde{W}_i(SI_i)^2 \end{aligned} \tag{4.32}$$

where $\tilde{V}_i = V_i^* - V_i$ and $\tilde{W}_i = \tilde{V}_i + W_i$.

$$\begin{aligned} W_i(k+1) &= W_i(k) + \Delta W_i(k) \\ \tilde{W}_i(k+1) &= \tilde{V}_i(k+1) + W_i(k+1) \\ &= \tilde{V}_i(k) + W_i(k) + \Delta W_i(k) \\ &= \tilde{W}_i(k) - K_{wi}\tilde{W}_i(SI_i)^2 \\ &= \left[1 - K_{wi}(SI_i)^2\right]\tilde{W}_i(k) \end{aligned} \tag{4.33}$$

□

Theorem 5. *(Closed-loop Stability): Given the initial UAS state $x(0)$ and the BELBIC estimated weights of Amygdala ($V_l(0)$) and Orbitofrontal Cortex ($W_l(0)$) be bounded in the set $\mathbf{\Lambda}$. Let the BELBIC be tuned and estimated control policy be given as* (4.21), (4.22) *and* (4.18) *respectively. Then, there exists positive constants, K_v, K_w, satisfying Theorem 4 such that UAS state, $x(t)$ and BELBIC weights estimation errors are all asymptotically stable.*

Proof. Let's consider the u_s is a stable controller for the following system:

$$\dot{x} = f(x) + g(x)u_s \tag{4.34}$$

There is a Lyapunov function $L_s(x)$ which guarantees the stability of the whole system:

$$L_s(x) = \frac{1}{2}x^T x$$

Taking the first derivative, we have:

$$\begin{aligned} \dot{L}_s(x) &= x^T \dot{x} \\ &= x^T[f(x) + g(x)u_s] \\ &\leq -lx^T x \quad , \ l > 0 \end{aligned} \tag{4.35}$$

To provide the stability analysis of the actual system, let's consider the u_a is an actual controller for the following system:

$$\dot{x} = f(x) + g(x)u_a \tag{4.36}$$

where u_a is as follows:

$$u_a = u_s + \tilde{u} \tag{4.37}$$

and $\tilde{u}$ is the controller which is given by the BELBIC model output MO. Considering the Lyapunov function $L_{MO}(x)$, the following is obtained:

$$L_{MO}(x) = A(\widetilde{MO})^2$$

Taking the first derivative, we have:

$$\begin{aligned}
\dot{L}_{MO}(x) &= A(\widetilde{MO})(\dot{\widetilde{MO}}) \\
&\leq -A(\widetilde{MO})^2 \\
&\leq -A(\tilde{u})^2
\end{aligned} \tag{4.38}$$

Considering the Lyapunov function $L_a(x)$, the stability proof of overall system is as follows:

$$L_a(x) = \frac{1}{2}x^T x$$

Taking the first derivative, we have:

$$\begin{aligned}
\dot{L}_a(x) &= x^T \dot{x} \\
&= x^T[f(x) + g(x)u_a] \\
&= x^T[f(x) + g(x)(u_s + \tilde{u})] \\
&= x^T[f(x) + g(x)u_s + g(x)\tilde{u}] \\
&= x^T[f(x) + g(x)u_s] + x^T g(x)\tilde{u} \\
&\leq -lx^T x + x^T g(x)\tilde{u} \\
&\leq -lx^T x + \frac{l}{2}x^T x + \frac{2}{l}(g(x)\tilde{u})^2 \\
&\leq -\frac{l}{2}x^T x + \frac{2}{l}(g(x)\tilde{u})^2 \\
&\leq -\frac{l}{2}x^T x - A(\tilde{u})^2
\end{aligned} \tag{4.39}$$

□

Remark 8. *Based on the BELBIC theory [41] and* (4.18)*, the optimal flocking control of MAS can be obtained while the estimated weights of Amygdala (V_i) and Orbitofrontal Cortex (W_i) are converging to desired targets. According to Theorem 4, estimated weights converge to desired targets asymptotically. Therefore, the designed R-BELBIC input U_i^{R-BEL}* (4.18) *converges to optimal flocking control of MAS asymptotically.*

In summary, the R-BELBIC-inspired flocking methodology for MAS is presented as a pseudo-code in **Algorithm** 4.

Algorithm 4 : The R-BELBIC-based flocking for MAS.

Initialization:
Set $V_i = 0$, $W_i = 0$, and $V_{th} = 0$, for $i = 1, ..., n$.
Define $ES_i =$ Objective function, for $i = 1, ..., n$.
for each iteration $t = t_s$ **do**
for each agent i **do**
Compute $SI_i = K^{\alpha}_{SI,i} u^{\alpha}_i + K^{\gamma}_{SI,i} u^{\gamma}_i + K^{CON}_{SI,i} \Upsilon_i (u_i, q_i, p_i)$
Compute $ES_i = K^{\alpha}_{ES,i} u^{\alpha}_i + K^{\gamma}_{ES,i} u^{\gamma}_i + K^{CON}_{ES,i} \Upsilon_i (u_i, q_i, p_i)$
Compute $A_i = V_i SI_i$
Compute $OC_i = W_i SI_i$
Compute $A_{th} = V_{th} \max (SI_i)$
Compute $BEL_i = \sum_i A_i - \sum_i OC_i$
Update V_i
Update W_i
Update V_{th}
end for
end for

In general, computing the computational complexity of an algorithm, we are interested in the maximum growth of the algorithm running time as the size of the problem gets larger. For these reasons, the $O(\cdot)$ notation is introduced [105].

Definition 1. *For two functions $\Psi(k)$ and $\Xi(k)$ of a non-negative parameter k, $\Psi(k) = O(\Xi(k))$ if there is a constant $\phi > 0$ such that for all sufficiently large k, $\Psi(k) \leq \phi \Xi(k)$. Therefore, the function $\phi \Xi(k)$ is an asymptotic upper bound on Ψ.*

Definition 2. *Polynomial-time algorithm (i.e., runs in polynomial time) if its running*

time $\Psi(k) = O(\Pi(k))$*, where* $\Pi(k)$ *is a polynomial function of the input size. Formally, polynomial-time algorithms are considered to be efficient, and problems for which polynomial-time algorithms exist are considered computationally efficient.*

Computing the computational complexity of the algorithm 4, it is on the order of $O(n)$, where n is the number of the agents. Comparing with other existing learning based algorithms, the computational complexity of the single output Multilayer Perceptron (MLP) is on the order of $O(cn)$, where c is the number of hidden neurons, and it is exponential for both Adaptive Neuro-Fuzzy Inference System (ANFIS), and Locally Linear Neuro-Fuzzy (LLNF) [106]. Therefore, the proposed method has lower computational complexity in comparison with other methods and is more appropriate for implementation in real-time applications.

4.4 Simulation Results

This section presents simulation results demonstrating the performance of the R-BELBIC for flocking of MAS in two different scenarios: 2-dimensional obstacle-free environment (Subsection 4.4.1) and 3-dimensional obstacle-free environments (Subsection 4.4.2). All simulations are carried out on a platform with following specifications: Windows Server 2012 R2 standard, Processor: Intel(R) Xeon(R) CPU E5-2680 0 @ 2.70GHz (4 processors), RAM: 8.00 GB.

In 2-dimensional scenario, a total of 150 agents were employed, while a total of 10 agents were employed in 3-dimensional scenario. Initial velocities are equal to zero, and positions randomly distributed in a squared area. The following parameters are used through all the simulations: $r = 1.2d_\alpha$ and $d_\alpha = 4$. For the σ-norm

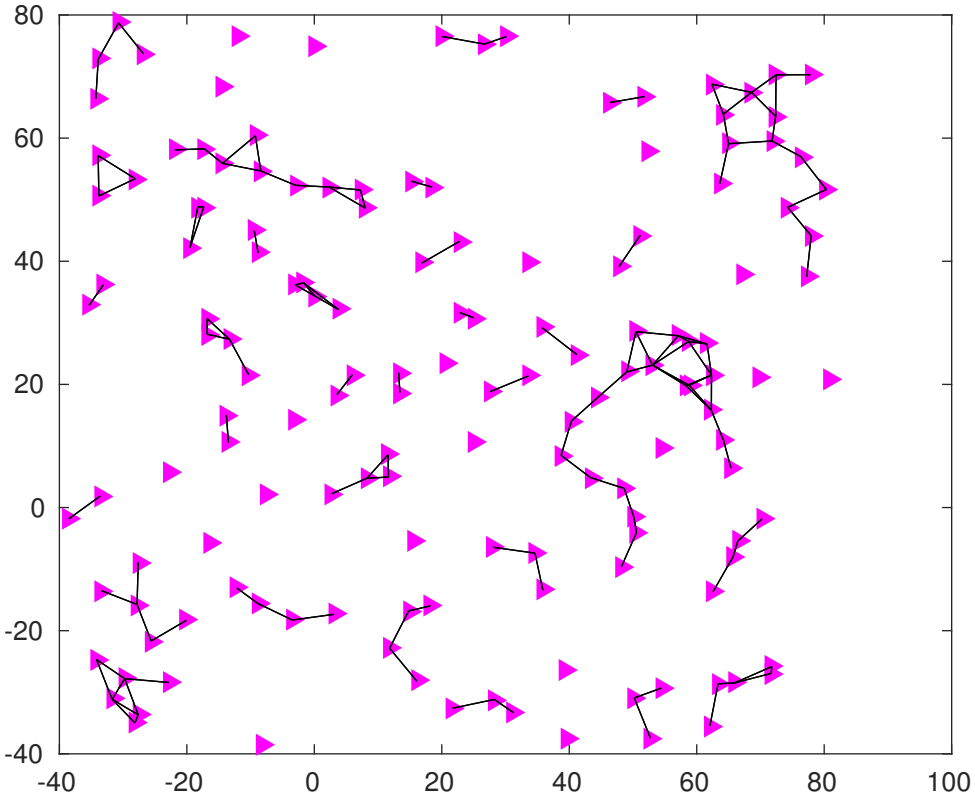

Figure 4.3: R-BELBIC-based Flocking of MAS. Simulation in an obstacle-free environment. 150 agents randomly distributed in a squared area at $t = 0s$.

the parameter $\epsilon = 0.1$, for $\phi(z)$ the parameters $a = b = 5$, for the bump function $\phi_\alpha(z)$ we use $h = 0.2$. The remaining parameters of the algorithm are specified in each subsection, as needed. In addition, all the design parameters are provided in Table 4.1.

Table 4.1: Parameters for designing SI and ES.

Parameter	Nominal Value	Parameter	Nominal Value
K_v	0.02	$K_{SI,i}^{CON}$	0.1
K_w	0.08	$K_{ES,i}^{\alpha}$	0.1
$K_{SI,i}^{\alpha}$	0.1	$K_{ES,i}^{\gamma}$	0.1
$K_{SI,i}^{\gamma}$	0.1	$K_{ES,i}^{CON}$	3

In this study, the performance of the proposed strategy when facing both un-

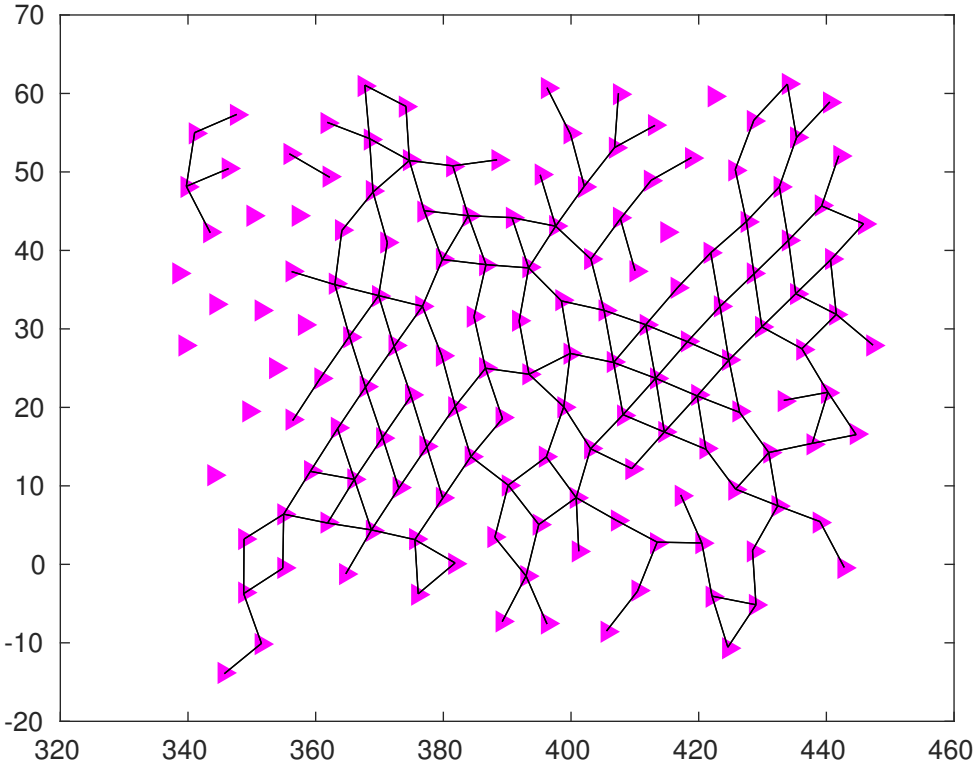

Figure 4.4: R-BELBIC-based Flocking of MAS. Simulation in an obstacle-free environment. At $t = 70s$ the 150 agents are flocking and have successfully formed a connected network.

known disturbances and system uncertainties is investigated. The disturbance appears in the interval between 10 to 20 second which affects all agents speed randomly. A random matrix of ± 10 percent is considered for applying uncertainties to the system.

4.4.1 Resilient Flocking in a 2-Dimensional Obstacles-Free Environment

Figure 4.3 and Figure 4.4 show two snapshots of the simulation in the obstacle-free environment. Figure 4.3 shows the 150 agents in their initial positions at $t = 0s$. It can be seen that the randomly distributed agents start forming an α-lattice. Figure 4.4 shows the agents at $t = 70s$ where they are flocking and have successfully formed a connected network. For comparison purposes, two similar experiments

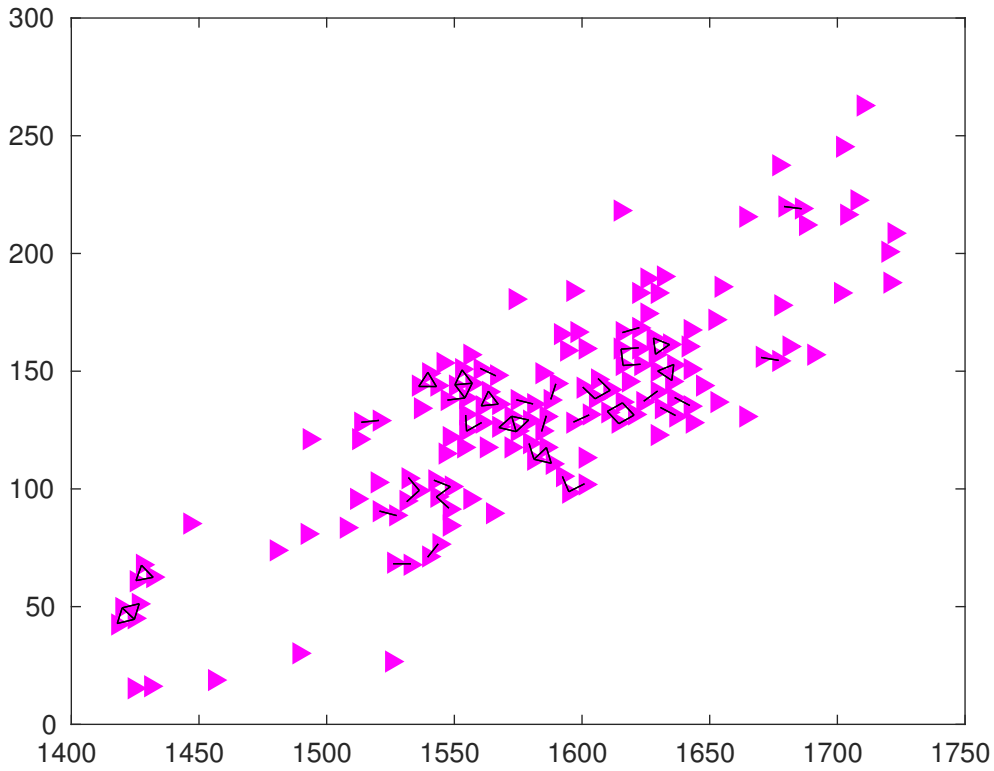

Figure 4.5: Flocking of MAS in [1]. Simulation in an obstacle-free environment. At $t = 70s$ the 150 agents are unable to successfully form a connected network.

were performed, but using the flocking algorithm introduced in [1] and [2] instead

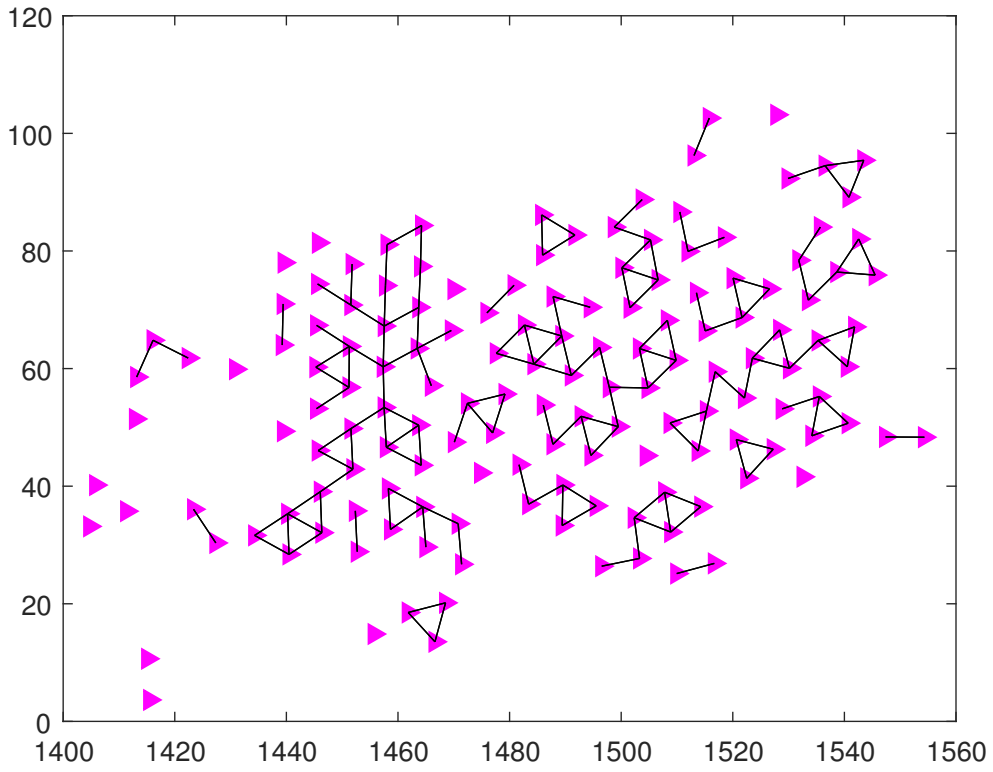

Figure 4.6: Flocking of MAS in [2]. Simulation in an obstacle-free environment. At $t = 70s$ the 150 agents are unable to successfully form a connected network.

of the R-BELBIC-based flocking. Figures 4.5 and 4.6 show the agents at $t = 70s$ where they are unable to successfully formed a connected network using flocking algorithm introduced in [1] and [2] respectively.

Figure 4.7 plots the velocity of all the agents in X-axis, and Y-axis for all controllers. As it can clearly be seen, the R-BELBIC controller could handle system uncertainties and reject disturbances and force all agents to have an agreement on their speed.

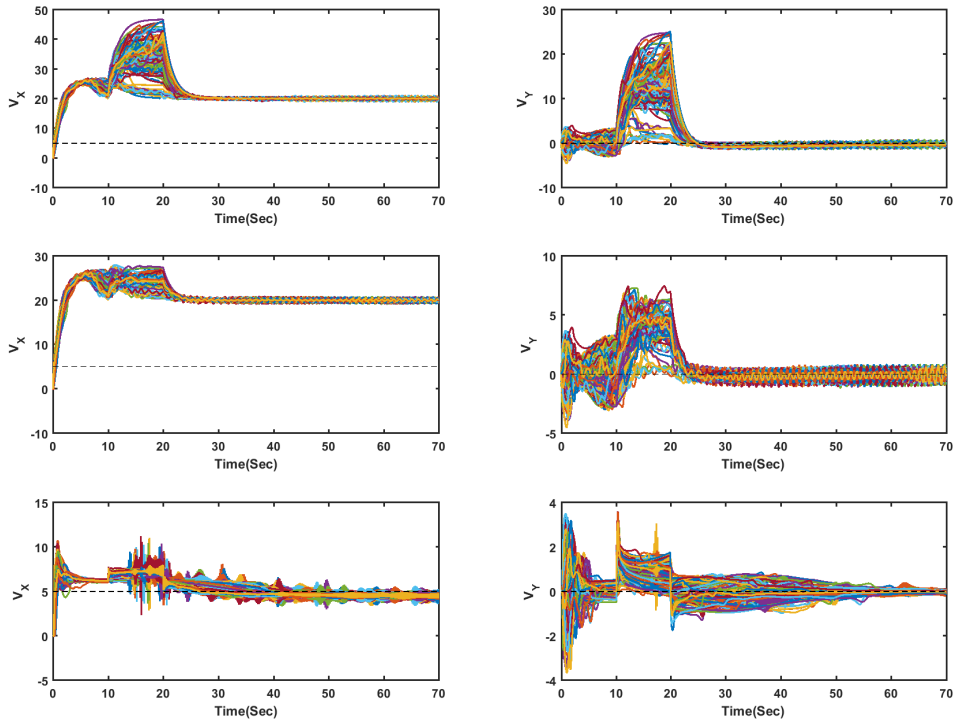

Figure 4.7: The velocity of all the agents in X-axis, and Y-axis, from left to right, respectively. The velocities of the conventional flocking controller [1] are plotted on the top row. The flocking algorithm in [2] in in middle row and the R-BELBIC-inspired control is in the bottom row. In all cases, the reference signal is shown in dashed black color. The disturbance appears in the interval between 10 to 20 second which affects all agents speed randomly.

4.4.2 Resilient Flocking in a 3-Dimensional Obstacle-Free Environment

Figure 4.8 and Figure 4.9 show two snapshots of the simulation in the obstacle-free environment. Figure 4.8 shows the 10 agents in their initial positions at $t = 0s$. It can be seen that the randomly distributed agents start forming an α-lattice. Figure 4.9 shows the agents at $t = 55s$ where they are flocking and have successfully formed a connected network.

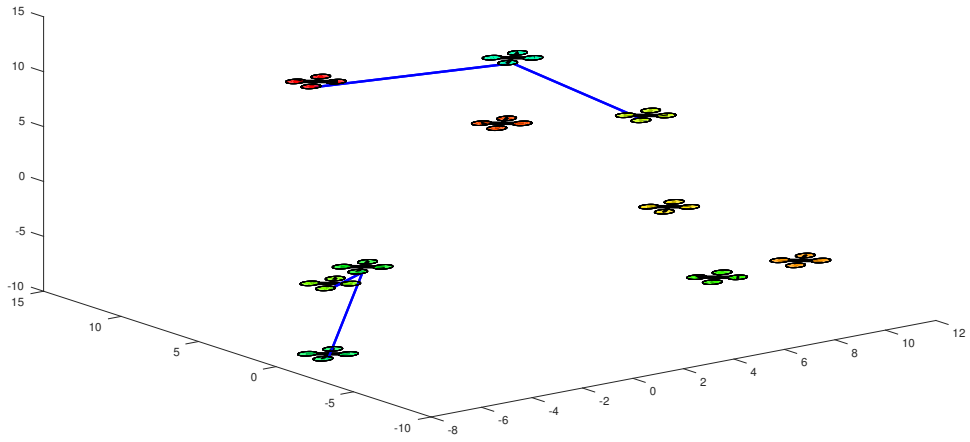

Figure 4.8: R-BELBIC-based Flocking of MAS. Simulation in an obstacle-free environment. 10 agents randomly distributed in a squared area at $t = 0s$.

For comparison purposes, two similar experiments were performed, but using the flocking algorithm introduced in [1] and [2] instead of the R-BELBIC-based flocking. Figures 4.10 and 4.11 show the agents at $t = 55s$ where they were unable to successfully form a connected network using the flocking algorithm introduced in [1] and [2] respectively.

Figure 4.12 plots the velocity of all the agents in X-axis, Y-axis, and Z-axis for all controllers. As it can clearly be seen, the R-BELBIC controller could handle system uncertainties and reject disturbances and force all agents to have an agreement on their speed.

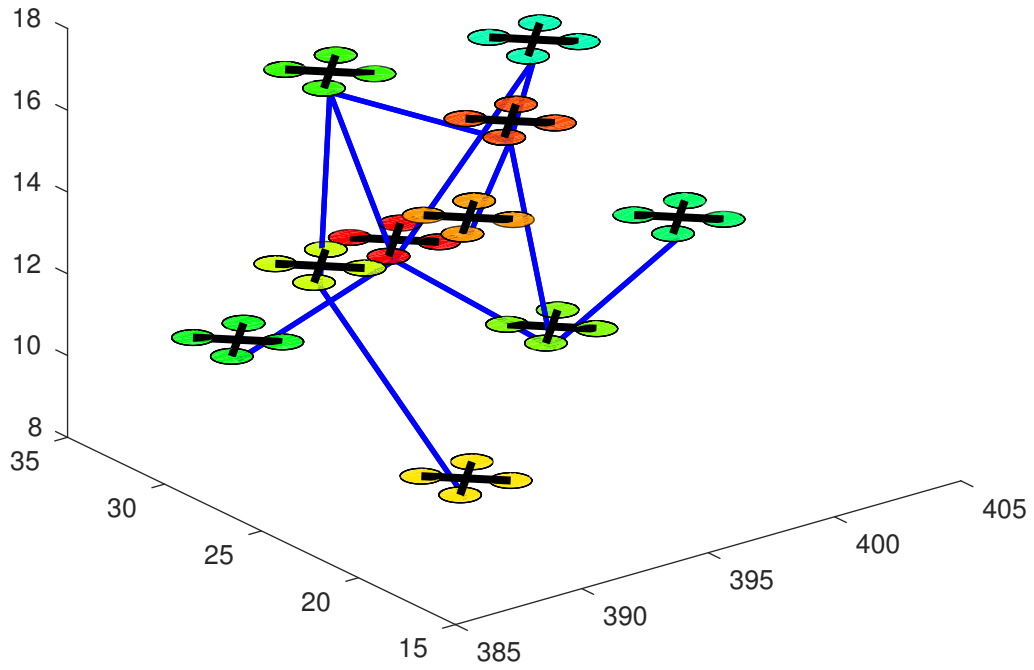

Figure 4.9: R-BELBIC-based Flocking of MAS. Simulation in an obstacle-free environment. At $t = 55s$ the 10 agents are flocking and have successfully formed a connected network.

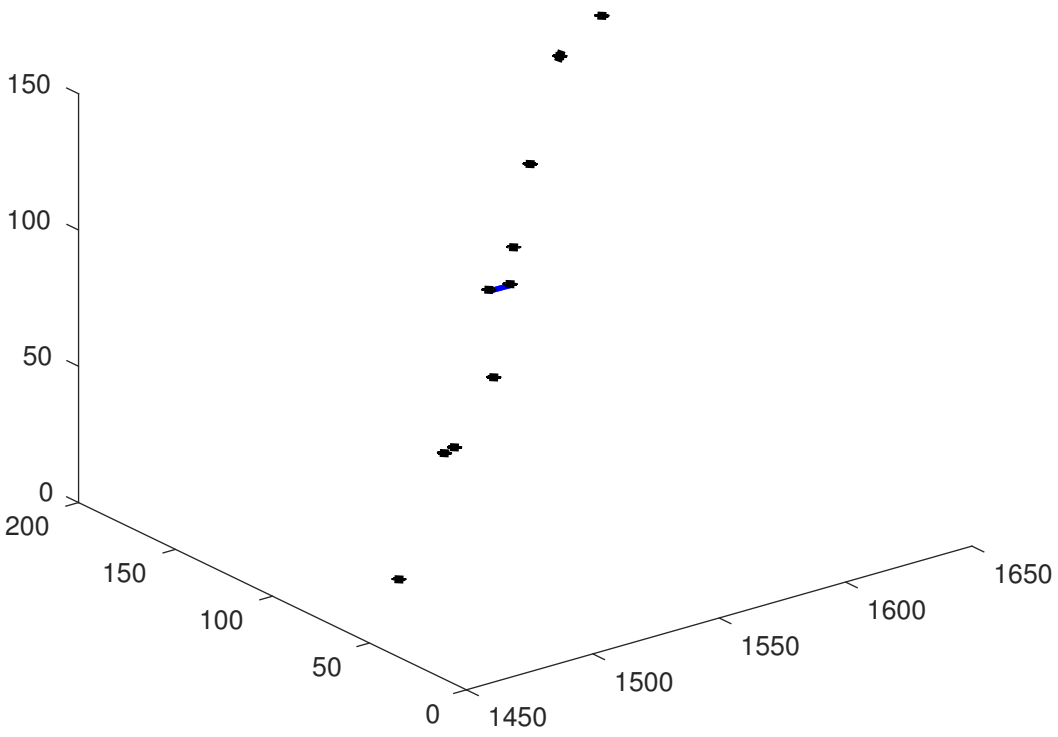

Figure 4.10: Flocking of MAS in [1]. Simulation in an obstacle-free environment. At $t = 55s$ the 10 agents were unable to successfully form a connected network.

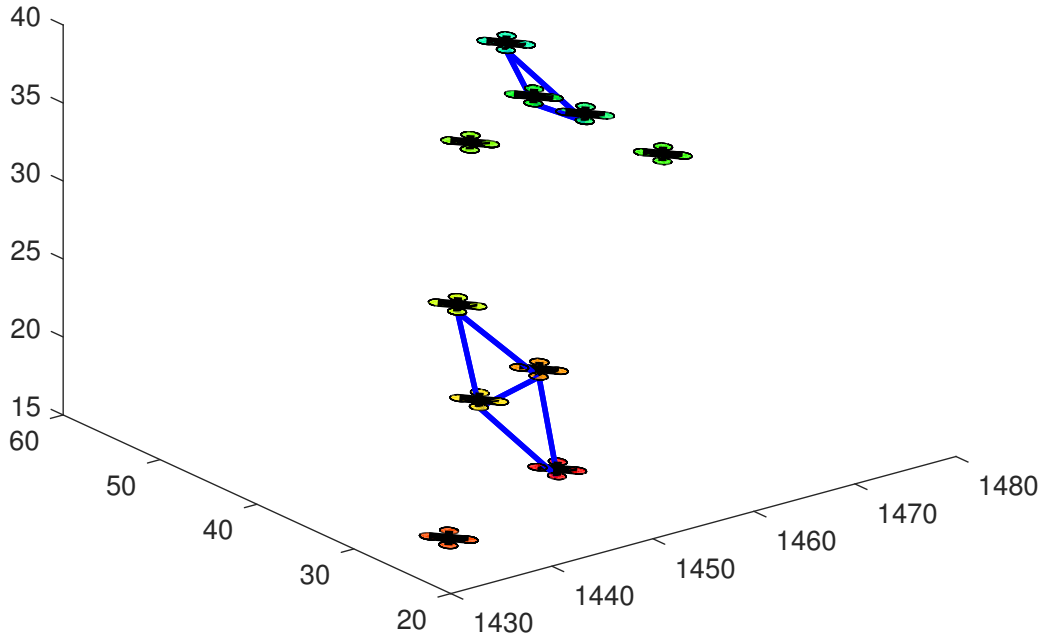

Figure 4.11: Flocking of MAS in [2]. Simulation in an obstacle-free environment. At $t = 55s$ the 10 agents were unable to successfully form a connected network.

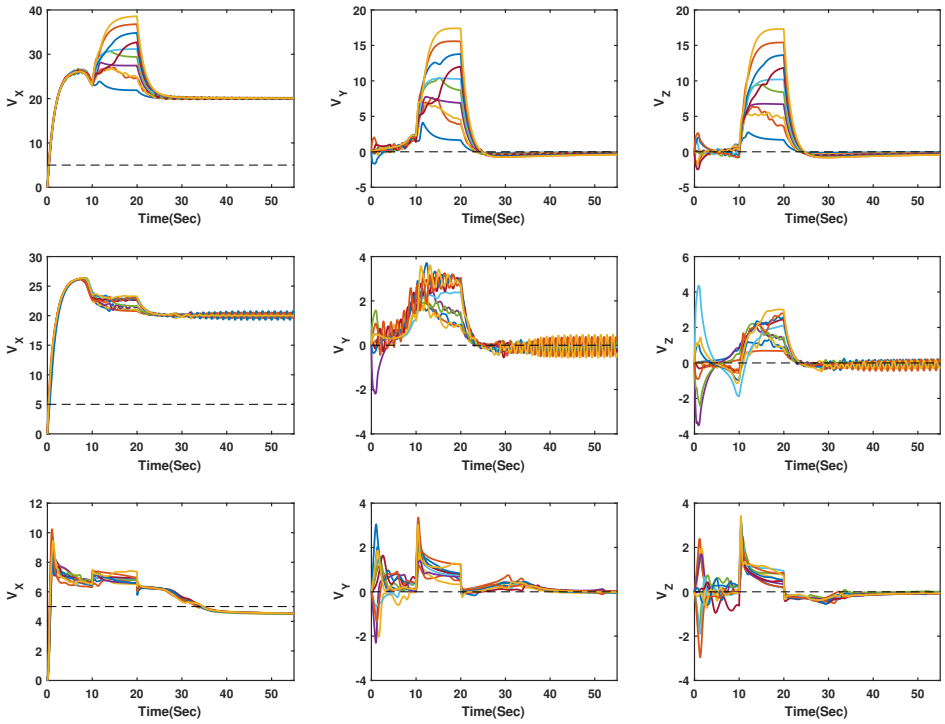

Figure 4.12: The velocity of all the agents in X-axis, Y-axis, and Z-axis, from left to right, respectively. The velocities of the conventional flocking controller [1] are plotted on the top row. The flocking algorithm in [2] in in middle row and the R-BELBIC-inspired control is in the bottom row. In all cases, the reference signal is shown in dashed black color. The disturbance appears in the interval between 10 to 20 second which affects all agents speed randomly.

4.5 Experimental Results

This section presents experimental results showing the performance of the R-BELBIC flocking control for MAS. Figure 4.13 shows the data flow of the proposed algorithm implementation setup.

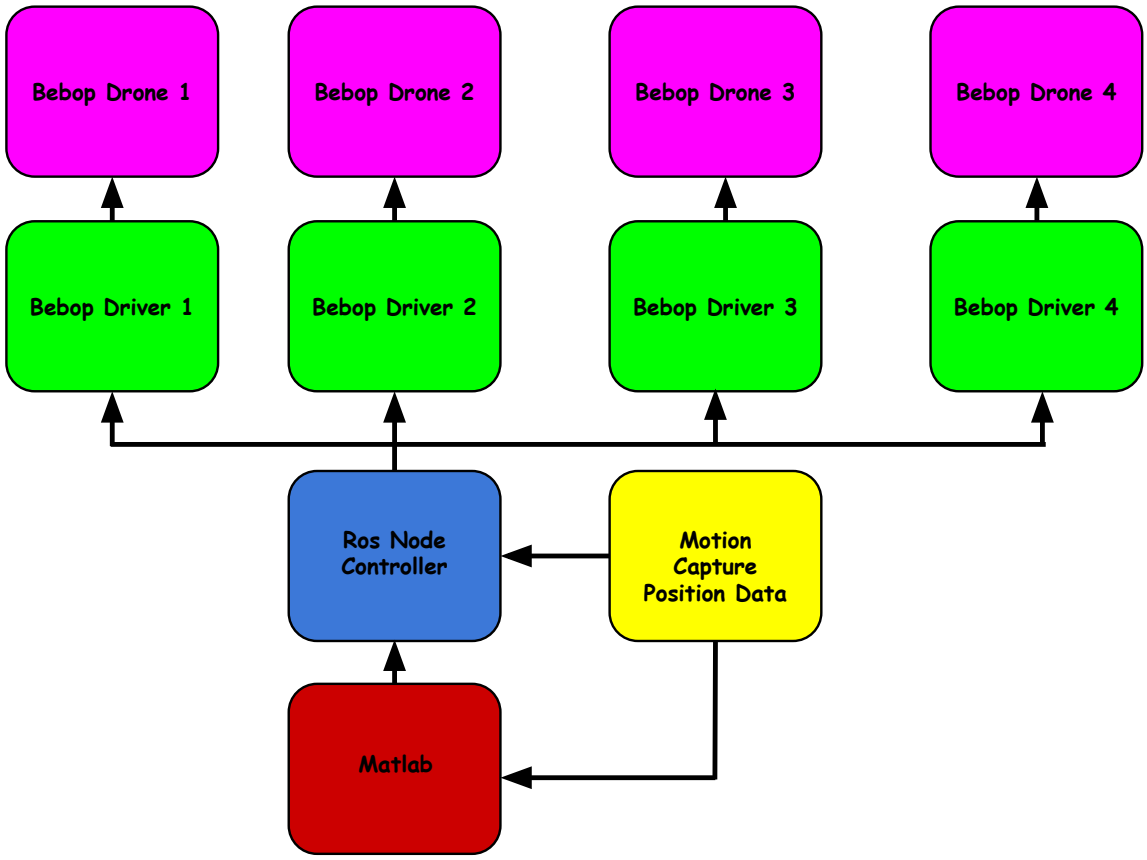

Figure 4.13: Data flow of the implementation setup using the R-BELBIC flocking algorithm.

Experimental Testbed

The platform implemented for validation of the proposed algorithm is available at the Unmanned Systems Laboratory (USL) from the University of Nevada - Reno,

see Figure 4.14. The Base Station of this testbed runs Ubuntu 14.04 OS and the Robot Operating System (ROS) environment in combination with Matlab. The proposed algorithms were coded in C/C++ and Matlab, and were executed at 100Hz. The aerial robots correspond to the Bebop drone, manufactured by Parrot.

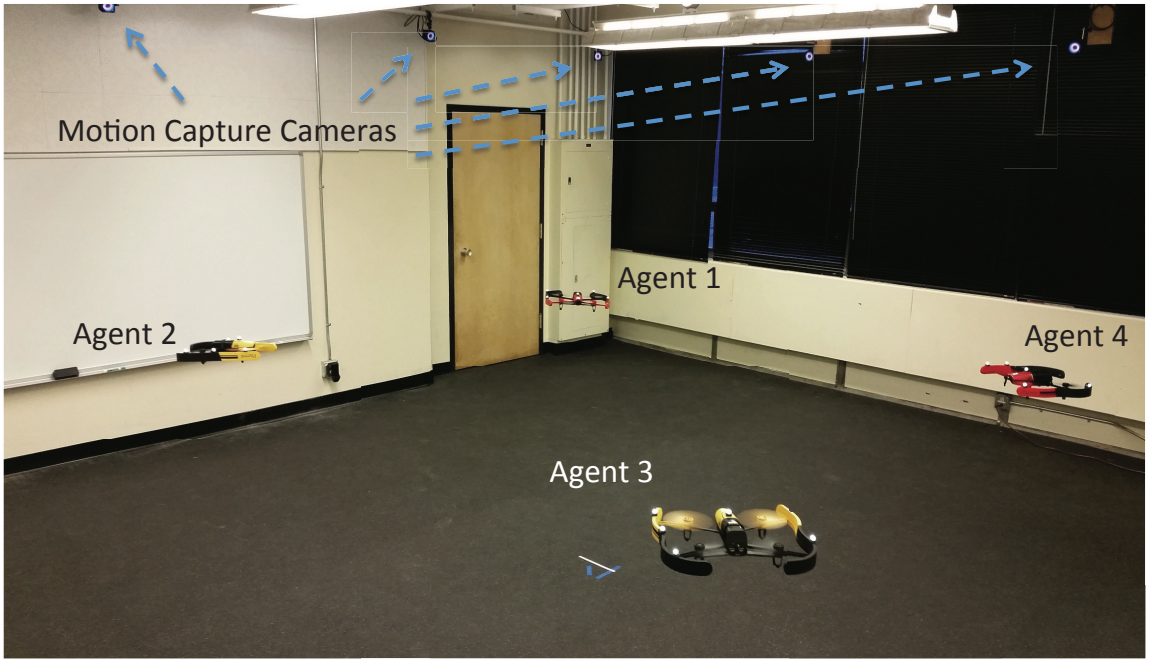

Figure 4.14: Experimental testbed: a set of four drones, a motion capture system, Ground Station computers, and WiFi links.

The 3-dimensional position of the UAS is obtained by means of a Motion Capture System (MCS) manufactured by OptiTrack. The information provided by the MCS is reported to the OptiTrack Interface PC by means of a Gigabyte Ethernet connection. Next, this information is sent to the Base Station PC by means of an Ethernet connection. The Base Station computer uses this information to execute the R-BELBIC algorithm and to calculate the control signals, which are then sent to the Bebop platforms by means of a WiFi link.

Real-time Experiments

The ultimate goal of this experimental application is to maintain a satisfactory flocking of the MAS, even when the model of the UAS is uncertain, and when unknown external factors affect the performance of the agents. In this experiment, a total of four UAS were employed, with initial velocities equal to zero, and positions randomly distributed in a squared area. The following parameters are used for the experiment: $r = 1.2d_\alpha$, while $d_\alpha = 2m$. For the σ-norm the parameter $\epsilon = 0.1$, for $\phi(z)$ the parameters $a = b = 5$, for the bump function $\phi_\alpha(z)$ we used $h = 0.2$.

Figure 4.15 plots the positions of the four UAS in (X, Y, Z) axis. It can be seen that all the agents successfully accomplishing the objective of our experimental application.

Figure 4.16 plots the velocities of the four UAS in (X, Y, Z) axis. It can be seen that all the agents agreed on the same speed, successfully accomplishing the objective of our experimental application.

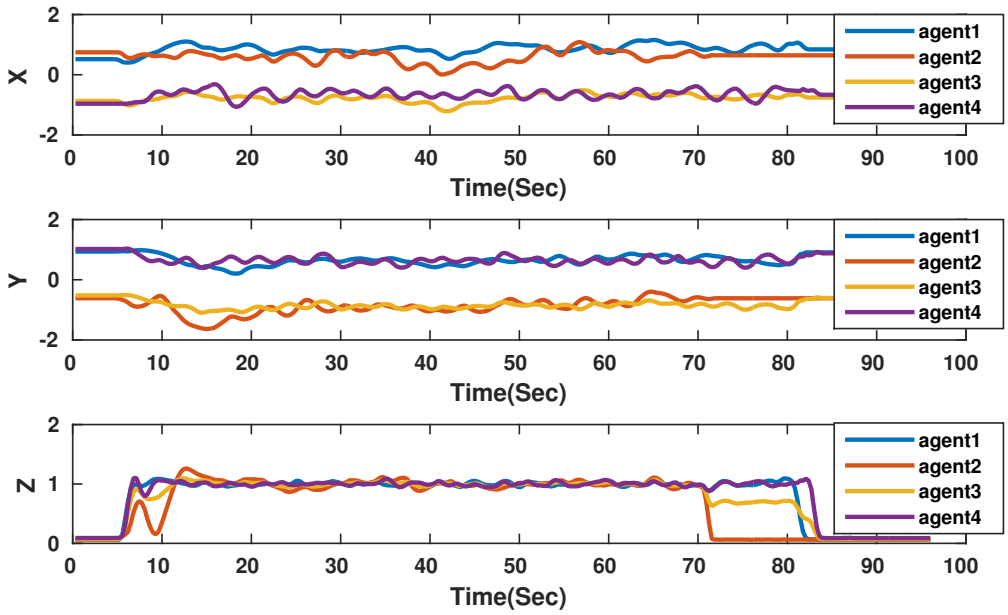

Figure 4.15: R-BELBIC flocking control of MAS: agreement in UAS positions in (X, Y, Z) axis.

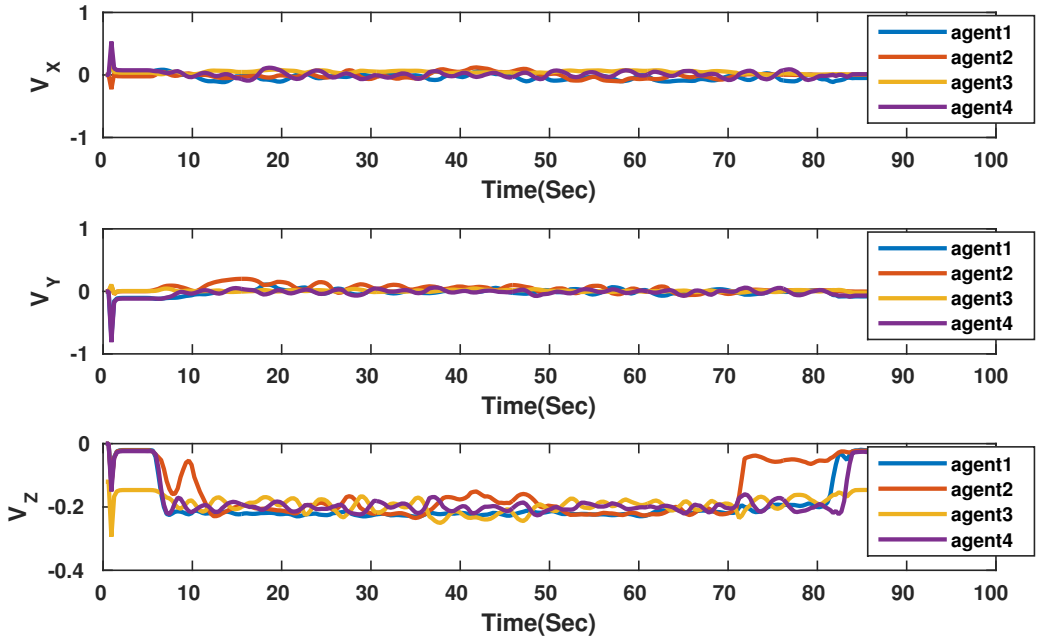

Figure 4.16: R-BELBIC flocking control of MAS: consensus in UAS velocities in (X, Y, Z) axis.

4.6 Conclusion

A neurobiologically-motivated intelligent distributed resilient controller based on a computational model of emotional learning in the mammalian limbic system was proposed for flocking control of MAS in presence of system uncertainties and unknown disturbances. The methodology, called R-BELBIC-inspired flocking, embeds a resilience mechanism with multi-objective properties into the flocking control strategy in a distributed manner. The results from both computer-aid simulation and experimental test demonstrate the effectiveness of proposed R-BELBIC based resilient distributed flocking control, as well as its applicability for real-time systems.

Chapter 5

A Game Theoretic Based Biologically-Inspired Distributed Intelligent Flocking Control for Multi-UAV Systems with Network Imperfections

Distributed controllers connected through the shared wireless communication network are necessary in designing distributed flocking control schemes for Networked multi-Unmanned Aerial Vehicles (UAVs). Network imperfections such as time delay, which has been considered as a challenging issue, commonly exists in the wireless network. Therefore, network imperfections should be taken

into account in designing distributed control algorithms for networked multi-UAV systems. Besides of network imperfections, the uncertainty from the complex environment and system dynamics is another critical challenge and cannot be ignored in advanced real-time control development. In this paper, a game theoretic based biologically-inspired distributed intelligent control methodology is proposed to overcome those challenges, i.e., networked imperfections and uncertainty from environment and system, in networked multi-UAV flocking. Considering the limited computational ability in the practical onboard micro-controller, the proposed method is adopted based on the game theory, and the emotional learning phenomenon in the mammalian limbic system. The learning capability and low computational complexity of the proposed technique makes it a propitious tool for implementing in networked multi-UAV flocking even in presence of the network imperfections and uncertainty from environment and system. Lyapunov analysis and computer-aid numerical simulation results of the implementation of the proposed methodology demonstrate the effectiveness of this algorithm.

5.1 Introduction

Motivation

Distributed coordination of networked multi-Unmanned Aerial Vehicles (UAVs) has been studied by diverse research communities in recent years. Due to the broad applications of *flocking* in real-world scenarios [1,7,34,107], most of the networked multi-UAV control methodologies are adopted based on the mathematical model of flocking [3,6,11]. In general, three basic rules (i.e., *separation*, *alignment*,

and *cohesion*) are considered for simulating the flocking behavior [7] which are observed in many living beings (i.e., birds, fish, bacteria, and insects) [15].

Improving the flocking behavior of networked multi-UAV systems has attracted several researchers in recent years [16,19,20,51]. Network imperfection, e.g., delay, commonly exists in communication due to the limited communication resource and heavy traffic in the network [108]. As a result, it is of paramount importance to address the challenges of network-induced delay and taking the influence of network-induced delays into account in designing control algorithms for multi-unmanned aerial vehicles. Besides of network imperfections, the uncertainty from the complex environment and system dynamics is another critical challenge and cannot be ignored in advanced applicable control development. Therefore, it is important to consider the uncertainties from environment and system in designing control algorithms.

Related works

Diverse research groups have attempted to address the issues arising from the effects of unknown disturbances and environment and system uncertainties in flocking control of multi-unmanned aerial vehicles/multi-agent systems. For example, the authors in [109] presented a distributed algorithms for the sensor networks by considering the effects of the imperfect communication such as link failures and channel noise. Closely related, a distributed control design for the discrete-time directed multi-agent systems with distributed network-induced delay has been proposed in [110]. Delay-independent flocking control of multi-agent systems have been addressed in [111], [112]. A distance constrained based adaptive flocking

control for the multi-agent system with network-induced delay was investigated in [113]. In [114], the authors studied the design of distributed formation recovery control for nonlinear heterogeneous multi-agent systems. Recently, coordinated control of two-wheel mobile robots with input network-induced delay was presented in [115].

Although all of these proposed methods perform well dealing with the effects of unknown disturbances and environment and system uncertainties, they still need the detailed information of the system. In this regard, the development of control strategies for dealing with the environment and system uncertainties with less dependency on the full knowledge of the system dynamics is of paramount importance.

In recent years, intelligent approaches have been extensively utilized for successfully solving diverse complex problems [23, 116]. Among them, biologically-inspired intelligent approaches have been received tremendous interests by many researchers because of their inherent potential to deal with computationally complex systems. *Emotional Learning* is one of such approaches, which takes advantage of a computational model of the amygdala in the mammalian limbic system [40]. This model, known as Brain Emotional Learning Based Intelligent Controller (BELBIC), consists of the *Amygdala*, *Orbitofrontal Cortex*, *Thalamus*, and *Sensory Input Cortex* as its main components. From a control systems point of view, BELBIC is a model-free controller (i.e., model dynamics are fully or partially unknown) which has shown promising performance under noise and system uncertainty [41].

Sensory Inputs (SI) and *Emotional Signal* (ES) are two main inputs to BELBIC model, and it is shown that the multi-objective problems could be solved by defining appropriate SI and ES [117, 118]. The flexibility in assigning different SI and

ES makes this controller a practical tool for implementation in real-time applications. Furthermore, BELBIC could effectively control a system even when the states of the system and the controller performance feedback are the only available information [41]. In addition, compared to other existing learning-based intelligent control methods, the computational complexity of BELBIC is on the order of $O(n)$ which makes it a suitable approach for real-time implementation.

Main Contributions

The main contribution of this paper is to develop a model-free distributed intelligent control methodology to overcome the challenges including the unknown disturbances and uncertainties from environment and system in networked multi-UAV systems. To this end, we propose a game theoretic based biologically-inspired distributed intelligent controller, which takes advantage of the game theory and the computational model of emotional learning in the mammalian limbic system. The proposed methodology has a low computational complexity which makes it a promising method for real-time applications. Furthermore, keeping the system complexity in a practically achievable limit, the proposed method delivers a controller with multi-objective properties (i.e., control effort optimization, handling the uncertainties from environment and system, and noise/disturbance rejection). Moreover, we provided the Lyapunov stability analysis to demonstrate that our proposed methodology guarantees the convergence of the designed control signals as well as maintain the system stability during the learning. The learning capability of the proposed approach is validated for flocking control of multi-unmanned aerial vehicles influenced by the unknown disturbances and uncertainties from environment and system with promising performance. Computer-based numerical

results of the implementation of the proposed methodology demonstrate the effectiveness of this algorithm for distributed intelligent flocking control of networked multi-UAV systems.

The rest of the paper is organized as follows. Section 5.2 presents the problem formulation and some preliminaries about flock modeling with network-induced delay, game theory, and emotional learning. Our main contribution is introduced in Section 5.3, which consists of a game theoretic based distributed intelligent flocking control strategy based on emotional learning. Section 5.4 presents numerical simulation results. The conclusion of the paper and future directions of our work are provided in Section 5.5.

5.2 Problem Formulation and Preliminaries

5.2.1 Flock Modelling

Assuming the movement of the flock in an m–dimensional space $(m = 2, 3)$, the equation of motion of the i^{th} agent with continuous-time double integrator dynamics could be described according to the following set of equations:

$$\begin{cases} \dot{q}_i(t) = p_i(t) \\ \dot{p}_i(t) = u_i(t) \quad , i = 1, 2, ..., n \end{cases} \tag{5.1}$$

where $u_i(t) \in \mathbb{R}^{\mathrm{m}}$ is the control input, $\{q_i(t), p_i(t)\} \in \mathbb{R}^{\mathrm{m}}$ are position, and velocity of the i^{th} agent, respectively. Consider a dynamic graph $\mathcal{G}(\upsilon, \varepsilon(t))$ which consists of a set of vertices $\upsilon = \{1, 2, ..., n\}$, and edges $\varepsilon(t) \subseteq \{(i, j) : i, j \in \upsilon, j \neq i\}$.

Each vertex represents an agent of the flock while a communication link between a pair of agents is represented by an edge.

$N_i^\alpha(t) = \{j \in \upsilon_\alpha : \ \|q_j(t) - q_i(t)\| < r, j \neq i\}$ is the neighborhood set of agent i, where the range of interaction between agents i and agent j is defined by a positive constant r, and $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^{\mathrm{m}}$. Solving the set of algebraic conditions: $\|q_j(t) - q_i(t)\| = d \quad \forall j \in N_i^\alpha(t)$, we could describe the geometric model of the flock, i.e., the α-lattice [1], where the distance between two neighbors i and j is represented by a positive constant d.

To avoid the singularity of the collective potential function at $q_i(t) = q_j(t)$, the σ-norm (i.e., $\|\cdot\|_\sigma$) is defined where $\|z\|_\sigma = \frac{1}{\epsilon}[\sqrt{1 + \epsilon\|z\|^2} - 1]$, and ϵ is a positive constant. To resolve the singularity problem, the set of algebraic conditions can be rewritten as: $\|q_j(t) - q_i(t)\|_\sigma = d_\alpha \quad \forall j \in N_i^\alpha(t)$.

A smooth collective potential function $V(q) = \frac{1}{2}\sum_i \sum_{j \neq i} \psi_\alpha(\|q_j(t) - q_i(t)\|_\sigma)$ can be obtained by considering the above-mentioned constraints, where $\psi_\alpha(z) = \int_{d_\alpha}^{z} \phi_\alpha(s)ds$ is a smooth pairwise potential function with $\phi_\alpha(z) = \rho_h(z/r_\alpha)\phi(z - d_\alpha)$, $\phi(z) = \frac{1}{2}[(a+b)\sigma_1(z+c) + (a-b)]$, and $\sigma_1(z) = z/\sqrt{1+z^2}$.

A possible choice for defining $\rho(z)$, which is a scalar bump function that smoothly varies between [0,1], is [1]:

$$\begin{cases} 1, & z \in [0, h) \\ \frac{1}{2}\left[1 + \cos\left(\pi \frac{(z-h)}{(1-h)}\right)\right], & z \in [h, 1] \\ 0, & \text{otherwise} \end{cases} \tag{5.2}$$

$u_i(t) = u_i^\alpha + u_i^\gamma$ is the flocking control algorithm introduced in [1], which consists of two main terms:

(i). u_i^α is the interaction component between two α-agents and is defined as follows:

$$u_i^\alpha = c_1^\alpha \sum_{j \in N_i^\alpha} \phi_\alpha(\|q_j(t) - q_i(t)\|_\sigma)\mathbf{n}_{i,j} + c_2^\alpha \sum_{j \in N_i^\alpha} a_{ij}(q)(p_j(t) - p_i(t)) \tag{5.3}$$

where c_1^α and c_2^α are positive constants. The terms $\mathbf{n}_{i,j}$ and $a_{ij}(q)$ are vector and the elements of the spatial adjacency matrix $A(q)$, respectively, which are described as follows:

$$\mathbf{n}_{i,j} = \frac{q_j(t) - q_i(t)}{\sqrt{1 + \epsilon\|q_j(t) - q_i(t)\|^2}}$$

$$a_{ij}(q) = \rho_h(\|q_j(t) - q_i(t)\|_\sigma / r_\alpha) \in [0, 1], j \neq i$$

where, $r_\alpha = \|r\|_\sigma$, and $a_{ii}(q) = 0$ for all i and q.

(ii). u_i^γ is a goal component which consists of a distributed navigational feedback term and is defined as follows:

$$u_i^\gamma = -c_1^\gamma \sigma_1(q_i - q_r) - c_2^\gamma(p_i - p_r) \tag{5.4}$$

where c_1^γ and c_2^γ are positive constants.

5.2.2 Network-induced Delays

Assuming that the state of agent i gets to agent j after passing through a communication channel with network-induced delay $\tau_{ij} > 0$, the u_i^α could be rewritten as:

$$u_i^\alpha = c_1^\alpha \sum_{j \in N_i^\alpha} \phi_\alpha(\|q_j(t-\tau_{ij}) - q_i(t-\tau_{ij})\|_\sigma)\mathbf{n}_{i,j}$$
$$+ c_2^\alpha \sum_{j \in N_i^\alpha} a_{ij}(q)(p_j(t-\tau_{ij}) - p_i(t-\tau_{ij})) \quad (5.5)$$

In this paper, we consider the case where the network-induced delays in all channels are equal to $\tau > 0$. Although the delay is deterministic, it is unknown. The proposed method can effectively handle this unknown delay.

5.2.3 Game Theory and Control

Generally, the following three elements are needed to completely describe a game.

(i). Set of players Π, $\Pi = \{1,\ 2,\ ...,\ M\}$ and M is the number of players.

(ii). A set of strategies ($\mathcal{S}_\pi$) for each player $\pi \in \Pi$. Then, the joint strategy set is $\mathcal{S} = \mathcal{S}_1 \times ... \times \mathcal{S}_\pi$. A joint strategy could be described as $s = (s_1,\ s_2,\ ...,\ s_M)$, $s \in \mathcal{S}$.

(iii). A cost function ($\mathcal{C}_\pi : \mathcal{S} \rightarrow R$) for each player $\pi \in \Pi$. Player π prefers the joint strategy s to $s^{'}$ if and only if the corresponding cost function to s is smaller than the corresponding cost function to $s^{'}$ (i.e., $\mathcal{C}_\pi(s) < \mathcal{C}_\pi(s^{'})$).

There are several classes of games, in which team games has a great relationship with control theory, among others [119]. Based on the flocking model of networked multi-UAV described in Section 5.2.1, each agent is only sensing its local information and its own cost function. Therefore, this flocking problem could be considered as a non-cooperative differential game problem [120, 121].

The joint strategy s^* is a Nash equilibrium if the following inequality holds.

$$\mathcal{C}_i(s_1^*, ..., s_\pi^*, ..., s_M^*) \leq \mathcal{C}_i(s_1^*, ..., s_\pi^{'}, ..., s_M^*)$$
$$\forall s^{'} \in \mathcal{S}_\pi \ , \ \forall \pi \in \Pi. \tag{5.6}$$

5.2.4 Brain Emotional Learning-Based Intelligent Controller

Brain Emotional Learning Based Intelligent Controller (BELBIC) is one of the neurobiologically-motivated intelligent methodologies, which is based on the computational model of emotional learning observed in the mammalian limbic system proposed in [40]. This model (depicted in Figure 5.1), has two main parts: *Amygdala*, and *Orbitofrontal Cortex*. Amygdala is responsible for immediate learning, while Orbitofrontal Cortex is responsible for inhibition of any inappropriate learning happening in the Amygdala. *Sensory Inputs* (SI) and *Emotional Signal* (ES) are two main inputs to the BELBIC model.

The output of the BELBIC model (MO) can be defined as

$$MO = \sum_l A_l - \sum_l OC_l \tag{5.7}$$

which is calculated by the difference between the Amygdala outputs (A_l) and the Orbitofrontal Cortex outputs (OC_l). Here, l is the number of sensory inputs.

The Orbitofrontal Cortex and the Amygdala outputs are calculated by the summation of all their corresponding nodes, where the output of each node is described as:

$$A_l = V_l SI_l \tag{5.8}$$

$$OC_l = W_l SI_l \tag{5.9}$$

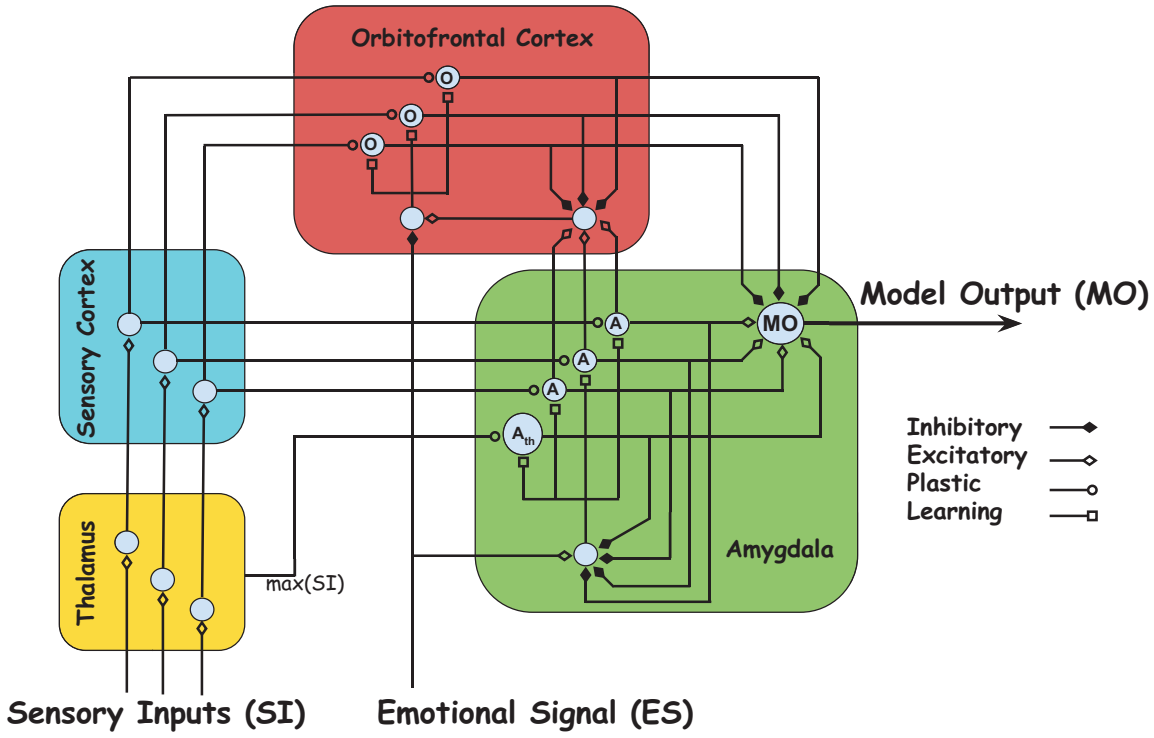

Figure 5.1: Computational model of emotional learning.

where SI_l is the l^{th} sensory input, V_l is the weight of the Amygdala, and W_l is the weight of the Orbitofrontal Cortex. The following equations are employed for updating V_l and W_l, respectively:

$$\Delta V_l = K_v SI_l \max\left(0, ES - \sum_l A_l\right) \tag{5.10}$$

$$\Delta W_l = K_w SI_l \left(MO - ES\right) \tag{5.11}$$

where, K_w and K_v are the learning rates.

The maximum of all SIs is another input considered in the model. This signal (i.e., A_{th}), which is directly sent from the Thalamus to the Amygdala, is defined as:

$$A_{th} = V_{th} \max\left(SI_l\right) \tag{5.12}$$

where V_{th} is the weight and the corresponding update law is the same as Equation (5.10).

5.2.5 Objectives

Based on the flocking model of networked multi-UAV described in Section 5.2.1, and by leveraging the game theory, and computational model of emotional learning in the mammalian limbic system (i.e., BELBIC) introduced in Sections 5.2.3 and 5.2.4, the objective is to design control signals $u_i(t)$, $i = 1, ..., n$ and n is the number of UAVs, as a game theoretic based biologically-inspired distributed intelligent controller, for flocking control of multi-unmanned aerial vehicles. Specifically, the proposed game theoretic based distributed controller is designed to maintain the motion of all agents in the flock in the events of network-induced delay.

5.3 Distributed Intelligent Flocking Control of Networked multi-UAV Systems using Game Theoretic based Emotional Learning

5.3.1 System Design

Generally, the intelligent techniques could be utilized for solving different control problems via direct or indirect approaches. In the direct mode, the intelligent method is utilized as a controller block, while it is employed for obtaining the controller's parameters in the indirect mode. In this paper, we propose a game theoretic based biologically-inspired distributed intelligent control technique which is based on the computational model of emotional learning in the mammal's brain and utilized it as a controller (i.e., direct mode), for intelligent flocking control of

multi-unmanned aerial vehicles with network-induced delay.

Specifically, our focus is on the design of a game theoretic based biologically-inspired distributed intelligent controller for flocking of networked multi-UAV by using BELBIC because the implementation of it could be accomplished without increasing the complexity of the overall system. The BELBIC architecture implemented in this work is shown in Figure 5.2. This figure demonstrates a closed loop configuration which consists of the following blocks: (i) BELBIC block, (ii) Sensory inputs (SI) function block, (iii) Emotional signal (ES) generator block, and finally (iv) a block for the plant. This architecture implicitly demonstrates the overall emotional learning based control concept, which consists of the action selection mechanism, the critic, and the learning algorithm [41].

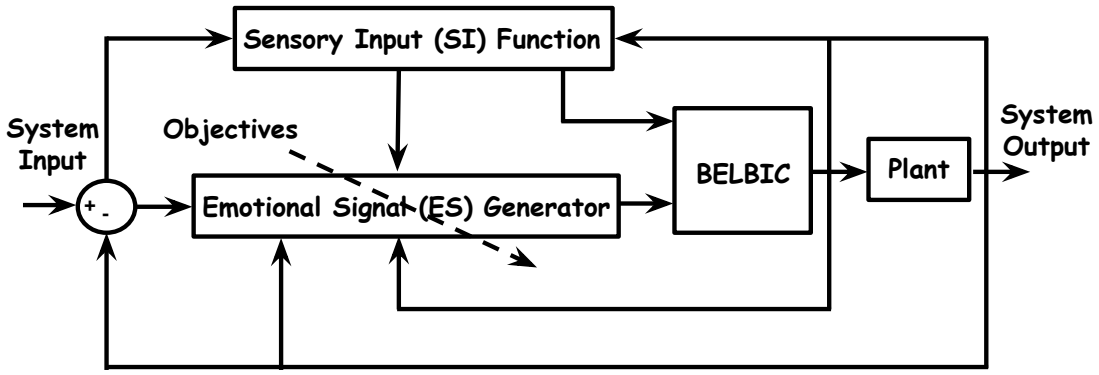

Figure 5.2: BELBIC in the control loop.

5.3.2 Emotional Signal and Sensory Input Development

Fundamentally, BELBIC is an action selection technique, in which action is produced based on the sensory input (SI) and the emotional signal (ES). The general

forms of SI and ES are given as follows:

$$SI = G\left(e, r, y, u\right) \tag{5.13}$$

$$ES = F\left(e, r, y, u\right) \tag{5.14}$$

where e is the system error, r is the system input, y is the system output, and u is the control effort.

The control objectives (e.g., reference tracking and optimal control) could implicitly be decided by choosing the adequate ES. For example, it is possible to choose the ES for achieving a better reference tracking performance, for reducing the overshoot, and/or for the energy expense minimization, among others.

Aiming at designing a game theoretic based distributed intelligent control for flocking control of multi-unmanned aerial vehicles, the proposed game theoretic based biologically-inspired technique will focus on improving: (i) reference tracking performance, (ii) network-induced delay handling, (iii) control effort optimization, and (iv) disturbance rejection.

To accomplish these objectives, for each of the control inputs (i.e., $\{u_1,...,u_n\}$), the SI_i will be designed as:

$$SI_i = K^{\alpha}_{SI,i} u^{\alpha}_i + K^{\gamma}_{SI,i} u^{\gamma}_i \tag{5.15}$$

The objective of the overall flock is to form an α–lattice over a finite time interval while tracking a target, which is equivalent to minimizing the following cost function by each individual agent:

$$\mathcal{C}_i = \frac{K^{\alpha}_{\mathcal{C},i}}{2} \parallel u^{\alpha}_i \parallel^2 + \frac{K^{\gamma}_{\mathcal{C},i}}{2} \parallel u^{\gamma}_i \parallel^2 \tag{5.16}$$

Minimizing the cost function means that every agent will try to minimize the total flocking error over a finite time interval, while at the same time minimizing its control effort, therefore the ES_i will be designed as:

$$ES_i = \mathcal{C}_i = K^{\alpha}_{ES,i} \parallel u^{\alpha}_i \parallel^2 + K^{\gamma}_{ES,i} \parallel u^{\gamma}_i \parallel^2 \tag{5.17}$$

where $i = \{1, ...n\}$, $K^{\alpha}_{SI,i}$, $K^{\gamma}_{SI,i}$, $K^{\alpha}_{ES,i} = \frac{K^{\alpha}_{\mathcal{C},i}}{2}$, and $K^{\gamma}_{ES,i} = \frac{K^{\gamma}_{\mathcal{C},i}}{2}$ are positive gains.

Solving the inequality in (5.6) for the above cost functions, the Nash strategies could be obtained.

It should be mentioned that, we designed the ES in such a way that the increase in reference tracking error will generate a *negative emotion* in the system, which is then taken as an evidence for the unsatisfactory performance of the system. Therefore, the proposed controller will behave in such a way that it will always minimize the negative emotion which leads to the satisfactory performance of the system.

5.3.3 Learning-based Intelligent Flocking Control

In flocking control of networked multi-UAV systems, multiple performance considerations have to be taken into account all at the same time, therefore, it is a very interesting case for using biologically-inspired learning-based multi-objective methodologies like BELBIC. Designing a model-free distributed intelligent control for flocking of networked multi-UAV by considering the network-induced delay, in addition to designing a suitable controller for real-time implementation, encourages us to take advantage of the computational model of emotional learning in the mammals' limbic system, i.e., BELBIC.

From equations (5.15)-(5.17), the BELBIC-inspired distributed intelligent flock-

ing control strategy for networked multi-UAV is defined as

$$u_i^{BEL} = \sum_i V_i \times SI_i - \sum_i W_i \times SI_i \tag{5.18}$$

here, $i = 1, ..., n$ makes reference to each control input and n is the number of UAVs. Considering the results obtained in Theorem 6 and by substituting the Emotional Signal with equation (5.17) the BELBIC model output of the distributed intelligent control for flocking of networked multi-UAV could be obtained as follows:

$$MO_i = ES_i = \mathcal{C}_i = K_{ES,i}^{\alpha} \parallel u_i^{\alpha}\parallel^2 + K_{ES,i}^{\gamma} \parallel u_i^{\gamma}\parallel^2 \tag{5.19}$$

which is clearly satisfies our goal of distributed intelligent flocking control.

5.3.4 Stability Analysis

Theorem 6 is presenting the convergence of the weights of the Amygdala (V_l) and the Orbitofrontal Cortex (W_l). Theorem 7 is providing the closed-loop stability of the proposed controller and Remark 9 explains how the proposed method converges to distributed intelligent flocking control of networked multi-UAV.

Theorem 6. *Given the BELBIC design as* (5.15)–(5.19)*, there exists the positive BELBIC tuning parameters, K_v, K_w satisfying*

I. $\left|\left[1 - K_v \left(SI_l\right)^2\right]\right| < 1$

II. $\left|\left[1 - K_w \left(SI_l\right)^2\right]\right| < 1$

such that the BELBIC estimated weights of the Amygdala (V_l) and the Orbitofrontal Cortex (W_l) converge to the desired targets asymptotically.

Proof. See **Appendix A** □

Theorem 7. *(Closed-loop Stability): Given the initial networked multi-UAV state $x(0)$ and the BELBIC estimated weights of the Amygdala ($V_l(0)$) and the Orbitofrontal Cortex ($W_l(0)$) be bounded in the set $\mathbf{\Lambda}$. Let the BELBIC be tuned and estimated control policy be given as* (5.21), (5.22) *and* (5.18) *respectively. Then, there exists positive constants, K_v, K_w, satisfying Theorem 6 such that networked multi-UAV state, $x(t)$ and BELBIC weights estimation errors are all asymptotically stable.*

Proof. See **Appendix B** □

Remark 9. *Based on the BELBIC theory [41] and* (5.18)*, the distributed intelligent flocking control of networked multi-UAV can be obtained while the estimated weights of the Amygdala (V_l) and the Orbitofrontal Cortex (W_l) are converging to desired targets. According to Theorem 6, estimated weights converge to desired targets asymptotically. Therefore, the designed BELBIC input U_l (i.e., Equation* (5.18)*) converges to distributed intelligent flocking control of networked multi-UAV asymptotically.*

5.4 Simulation Results

This section presents computer-based simulation results showing the performance of the proposed game theoretic based biologically-inspired distributed intelligent flocking control of multi-unmanned aerial vehicles in an obstacle-free environment. A total of 20 unmanned aerial vehicles where employed, with initial veloc-

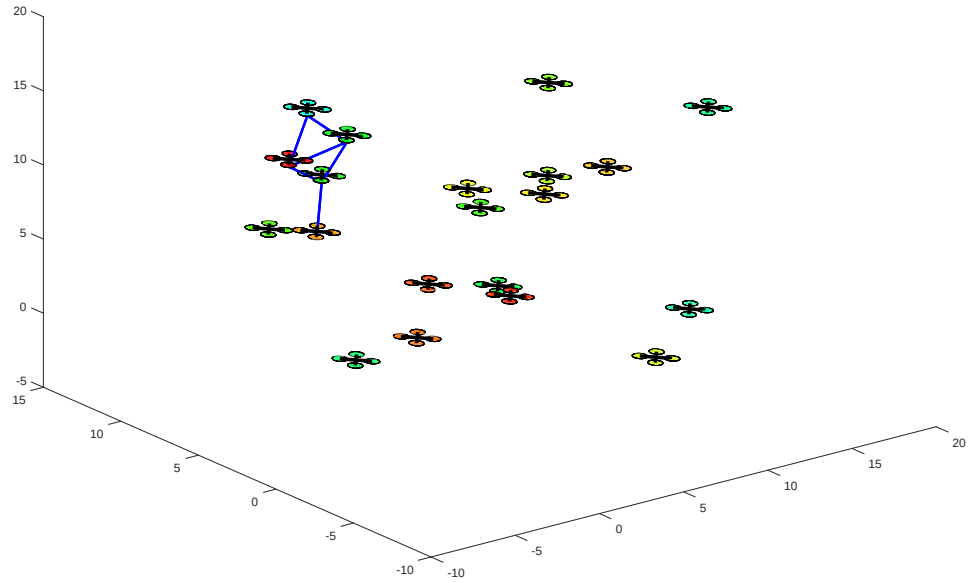

Figure 5.3: Simulation in an obstacle-free environment. 20 UAVs randomly distributed in a squared area at $t = 0s$.

ities equal to zero, and positions randomly distributed in a squared area. The following parameters are used through the simulation: $d_\alpha = 7$, $r = 1.2d_\alpha$, $d^{'} = 0.6d_\alpha$, $r^{'} = 1.2d^{'}$. For the σ-norm the parameter $\epsilon = 0.1$, for $\phi(z)$ the parameters $a = b = 5$, for the bump functions $\phi_\alpha(z)$ and $\phi_\beta(z)$, $h = 0.2$ and $h = 0.9$, respectively. The same network-induced delays (i.e., $\tau_{i,j} = \tau = 0.3$, $\forall i$) are considered for all unmanned aerial vehicles.

Figure 5.3 and Figure 5.4 show two snapshots of the simulation in the obstacle-free environment. Figure 5.3 shows the 50 UAVs in their initial positions at $t = 0s$ while Figure 5.4 shows the UAVs at $t = 40s$ where they are flocking and have successfully formed a connected network.

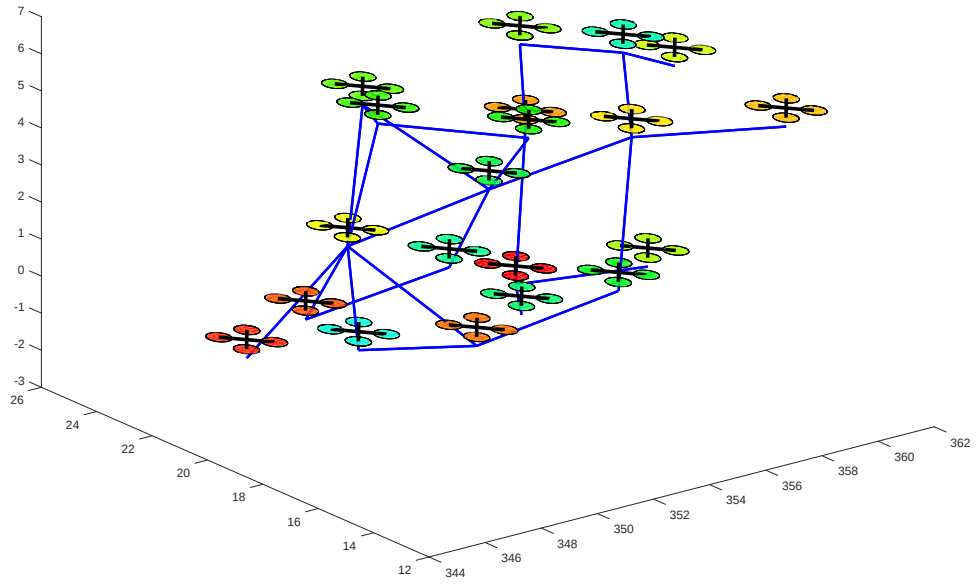

Figure 5.4: Simulation in an obstacle-free environment. At $t = 40s$ the 20 UAVs are flocking and have successfully formed a connected network.

For comparison purposes, a similar experiment was performed, but using the flocking algorithm introduced in [1] instead of the proposed algorithm. Figure 5.5 shows the velocities of all UAVs in X and Y axis for both methods in an obstacle-free environment under the influence of the networked-induced delay. The plot shows that, although the delay is deterministic, it is unknown and the proposed method could effectively handle the influences of the network-induced delay.

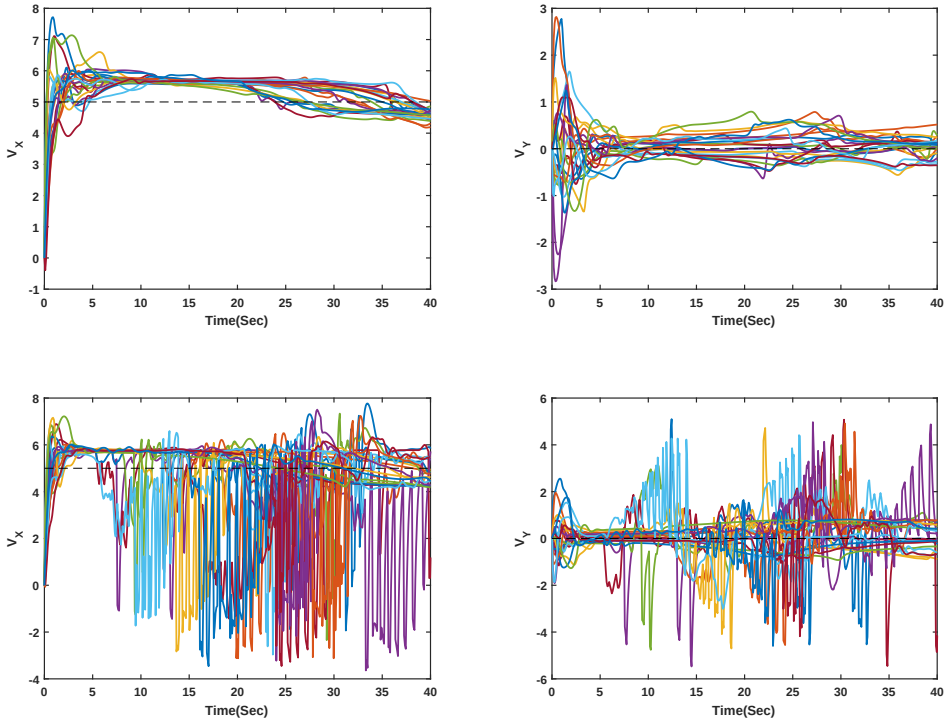

Figure 5.5: Velocities of all UAVs in X and Y axis for both methods in an obstacle-free environment under the influence of the networked-induced delay. The proposed method top row and the flocking algorithm proposed in [1] bottom row.

5.5 Conclusions

A game theoretic based biologically-inspired distributed intelligent control methodology is proposed to overcome the challenges of network-induced delay in flocking of networked multi-unmanned aerial vehicles. The methodology is adopted based on the emotional learning phenomenon in the mammalian limbic system. The learning capability and low computational complexity of the proposed technique makes it a promising tool for implementing in real-time networked multi-unmanned aerial vehicles flocking considering the influence of network-induced delay. Computer-based numerical results of the implementation of the proposed methodology demonstrate the effectiveness of this algorithm for distributed intelligent flocking control of networked multi-UAV.

5.6 Appendix A

5.6.1 Non-adapting Phase

Our goal is to investigate the output of the system in non-adapting phase (i.e., when the system completes its learning process) so the equations (5.10) and (5.11) which are the updating rules of Amygdala and Orbitofrontal Cortex, respectively, should be taken into consideration. In addition we make an assumption that the max function in equation (5.10) could be neglected. By substituting (5.8) and (5.9) in equation (5.7) the output of the model could be defined as follows:

$$\begin{aligned} MO =& \sum_l V_l SI_l - \sum_l W_l SI_l \\ =& \sum_l \left(V_l - W_l\right) SI_l \end{aligned} \tag{5.20}$$

$$\begin{aligned} \Delta V_l =& K_v SI_l \left(ES - \sum_l A_l\right) \\ =& K_v SI_l \left(ES - \sum_l V_l SI_l\right) \end{aligned} \tag{5.21}$$

$$\begin{aligned} \Delta W_l =& K_w SI_l \left(MO - ES\right) \\ =& K_w SI_l \left(\sum_l \left(V_l - W_l\right) SI_l - ES\right) \end{aligned} \tag{5.22}$$

When the learning process is completed (i.e., after system completes its learning process) the variations of the weights of Amygdala (ΔV_l) and Orbitofrontal Cortex (ΔW_l) will be equal to zero (i.e., $\Delta V_l = \Delta W_l = 0$). With the assumption of $SI_l \neq 0$

the following holds:

$$K_v SI_l \left(ES - \sum_l V_l SI_l \right) = 0$$
$$\Rightarrow \sum_l V_l SI_l = ES \tag{5.23}$$

$$K_w SI_l \left(\sum_l (V_l - W_l)\, SI_l - ES \right) = 0$$
$$\Rightarrow \sum_l (V_l - W_l)\, SI_l = ES$$
$$\Rightarrow \sum_l W_l SI_l = 0 \tag{5.24}$$

By substituting (5.23) and (5.24) in equation (5.20) the model output in non-adapting phase will be as follows:

$$MO = \sum_l (V_l - W_l)\, SI_l = \sum_l V_l SI_l = ES \tag{5.25}$$

5.6.2 Main Proof

Considering the results obtained in Subsection 5.6.1 the following should be achieved:

$$MO_l \rightarrow ES_l \tag{5.26}$$

Let's considering the V_l^* is the weight of Amygdala for each control input l when the system has been learned and let's $\hat{ES_l}$ be the Emotional Signal for each control input l during the adaptation phase. The following hold:

$$ES_l = V_l^* SI_l \;\; and \;\; \hat{ES_l} = V_l SI_l \tag{5.27}$$

$$\Delta V_l(k) = K_v SI_l \max\left(0, ES_l - \hat{ES_l}\right) \tag{5.28}$$

We will investigate the results of the following two cases:

I. $ES_l - \hat{ES}_l \geq 0$

II. $ES_l - \hat{ES}_l < 0$

Considering the case I., the proof can be achieved as follows:

$$\begin{aligned}
\Delta V_l(k) =& K_v SI_l \max\left(0, ES_l - \hat{ES}_l\right) \\
=& K_v SI_l \left(ES_l - \hat{ES}_l\right) \\
=& K_v SI_l \left(V_l^* SI_l - V_l SI_l\right) \\
=& K_v SI_l \left(V_l^* - V_l\right) SI_l \\
=& K_v SI_l \tilde{V}_l SI_l \\
=& K_v \tilde{V}_l \left(SI_l\right)^2
\end{aligned} \tag{5.29}$$

where $\tilde{V}_l = V_l^* - V_l$.

$$\begin{aligned}
V_l\left(k+1\right) =& V_l\left(k\right) + \Delta V_l(k) \\
\tilde{V}_l\left(k+1\right) =& V^* - V_l\left(k\right) - \Delta V_l(k) \\
=& \tilde{V}_l\left(k\right) - K_v \tilde{V}_l \left(SI_l\right)^2 \\
=& \left[1 - K_v \left(SI_l\right)^2\right] \tilde{V}_l\left(k\right)
\end{aligned} \tag{5.30}$$

Considering the case II., it is obvious that when $ES_l - \hat{ES}_l < 0$ the max function in equation (5.28) will force the adaptation in Amygdala to stop and the following hold:

$$\begin{aligned}
\Delta V_l(k) =& 0 \\
V_l\left(k+1\right) =& V_l\left(k\right) \\
\tilde{V}_l\left(k+1\right) =& \tilde{V}_l\left(k\right)
\end{aligned} \tag{5.31}$$

The proof can be achieved as follows:

$$\begin{aligned}\Delta W_l(k) =& K_w SI_l \left(MO_l - ES_l\right)\\ =& K_w SI_l \left(V_l SI_l - W_l SI_l - V_l^* SI_l\right)\\ =& K_w SI_l \left(-\left(V_l^* - V_l\right) SI_l - W_l SI_l\right)\\ =& K_w SI_l \left(\left(-\tilde{V}_l - W_l\right) SI_l\right)\\ =& - K_w \tilde{W}_l \left(SI_l\right)^2 \end{aligned} \tag{5.32}$$

where $\tilde{V}_l = V_l^* - V_l$ and $\tilde{W}_l = \tilde{V}_l + W_l$.

$$\begin{aligned} W_l\left(k+1\right) =& W_l\left(k\right) + \Delta W_l(k)\\ \tilde{W}_l\left(k+1\right) =& \tilde{V}_l\left(k+1\right) + W_l\left(k+1\right)\\ =& \tilde{V}_l\left(k\right) + W_l\left(k\right) + \Delta W_l(k)\\ =& \tilde{W}_l\left(k\right) - K_w \tilde{W}_l \left(SI_l\right)^2\\ =& \left[1 - K_w \left(SI_l\right)^2\right] \tilde{W}_l\left(k\right) \end{aligned} \tag{5.33}$$

5.7 Appendix B

Let's consider the u_s is a stable controller for the following system:

$$\dot{x} = f(x) + g(x)u_s \tag{5.34}$$

There is a Lyapunov function $L_s(x)$ which guarantees the stability of the whole system:

$$L_s(x) = \frac{1}{2} x^T x$$

Taking the first derivative, we have:

$$\begin{aligned}\dot{L}_s(x) &= x^T\dot{x} \\ &= x^T[f(x) + g(x)u_s] \\ &\leq -lx^Tx \quad , \ l > 0\end{aligned} \tag{5.35}$$

To provide the stability analysis of the actual system, let's consider the u_a is an actual controller for the following system:

$$\dot{x} = f(x) + g(x)u_a \tag{5.36}$$

where u_a is as follows:

$$u_a = u_s + \tilde{u} \tag{5.37}$$

and $\tilde{u}$ is the controller which is given by the BELBIC model output MO. Considering the Lyapunov function $L_{MO}(x)$, the following is obtained:

$$L_{MO}(x) = A(\widetilde{MO})^2$$

Taking the first derivative, we have:

$$\begin{aligned}\dot{L}_{MO}(x) &= A(\widetilde{MO})(\dot{\widetilde{MO}}) \\ &\leq -A(\widetilde{MO})^2 \\ &\leq -A(\tilde{u})^2\end{aligned} \tag{5.38}$$

Considering the Lyapunov function $L_a(x)$, the stability proof of overall system is as follows:

$$L_a(x) = \frac{1}{2}x^Tx$$

Taking the first derivative, we have:

$$\begin{aligned}
\dot{L}_a(x) &= x^T \dot{x} \\
&= x^T[f(x) + g(x)u_a] \\
&= x^T[f(x) + g(x)(u_s + \tilde{u})] \\
&= x^T[f(x) + g(x)u_s + g(x)\tilde{u}] \\
&= x^T[f(x) + g(x)u_s] + x^T g(x)\tilde{u} \\
&\leq -lx^T x + x^T g(x)\tilde{u} \\
&\leq -lx^T x + \frac{l}{2}x^T x + \frac{2}{l}(g(x)\tilde{u})^2 \\
&\leq -\frac{l}{2}x^T x + \frac{2}{l}(g(x)\tilde{u})^2 \\
&\leq -\frac{l}{2}x^T x - A(\tilde{u})^2
\end{aligned} \tag{5.39}$$

Chapter 6

Brain Emotional Learning-Based Path Planning and Intelligent Control Co-Design for Unmanned Aerial Vehicle in Presence of System Uncertainties and Dynamic Environment

This paper proposes a novel intelligent path planning and control co-design for Unmanned Aerial Vehicles (UAVs) in the presence of system uncertainties and dynamic environments. In order to simultaneously handle the uncertainties from

both the UAV platform itself and from the environment, a novel biologically-inspired approach based on a computational model of emotional learning in mammalian limbic system is adopted. The methodology, known as Brain Emotional Learning (BEL), is implemented in this application for the first time. Making use of the multi-objective properties and the real-time learning capabilities of BEL, the path planning and control co-design are applied in a synthetic UAV path planning scenario, successfully dealing with the challenges caused by system uncertainties and dynamic environments. A Lyapunov analysis demonstrates the convergence of the co-design, and a set of numerical results illustrate the effectiveness of the proposed approach. Furthermore, it is shown that the low computational complexity of the method guarantees its implementation in real-time applications.

6.1 Introduction

Path planning and control of Unmanned Aerial Vehicles (UAVs) are in demand for both military and civilian missions specifically in the mobile robotics area [17, 122, 123]. In general, the critical challenge in path planning is how to efficiently determine the UAV moving trajectory from current location to the desired location in a complex environment in real-time. Path planning strategies are also responsible for collision avoidance, obstacle avoidance, and goal satisfaction. Moreover, how to develop an intelligent and robust control to achieve the plan under an uncertain and complex environment is still a relevant and open problem. In recent years, both path planning and control of Unmanned Aerial Vehicles (UAVs) have been received a tremendous interests from robotics as well as control societies [17, 123].

Due to the strong interaction between path planning and real-time control, a

novel and effective co-design is required. Moreover, the path planning and control co-design in dynamic environment with unknown moving obstacle is a critical challenging problem that attracts enormous interests from many practical applications, e.g., outer space exploration, disaster search and rescue, etc. Diverse research efforts have been proposed aiming at addressing the issues arising from these conditions i.e., dynamic environments with unpredictable and unknown moving obstacles [123].

Related Works

In [124], the authors presented an evolutionary artificial potential field algorithm for dynamic path planning of mobile robot. Closely related, harmonic potential-based communication-aware navigation of mobile agents in cluttered spaces is proposed in [125]. Scalable lazy satisfiability modulo theory-based motion planning has been addressed in [126]. Moreover, authors in [127] studied a hierarchical path generation for distributed mission planning of UAVs. In [128], the authors introduced a path planning method with fractional potential fields for autonomous vehicles. Recently, a near-optimal decoupling principle for nonlinear stochastic systems arising in robotic path planning and control have been investigated in [129].

In general, most existing path planning researches available in the literature, see [122, 123], and the references therein, address three aspects, specifically, (i) the optimization of the planned path, (ii) the robustness and adaptivity for dealing with dynamic and uncertain environments, and (iii) the capability of dealing with unpredictable and unknown moving obstacles. However, these existing ap-

proaches are not designed for having multi-objective properties which is important for the path planning and control co-design due to the various constraints existing in UAVs. Furthermore, the computational complexity of these existing approaches is not suitable for real-time implementation.

In recent years, solving complex computational problems by means of biologically-inspired methods have been increasingly employed by many researchers [40, 41, 44]. Brain Emotional Learning (BEL) [41] is one of the most promising learning techniques. This methodology adopts the computational model developed in [40] to mimic the parts of the mammalian limbic system which are known for producing emotion, i.e., the Amygdala, Orbitofrontal Cortex, Thalamus, and Sensory Input Cortex. Strategies based on BEL have shown to be a very effective solution for improving the robustness as well as for handling uncertainties from both physical systems and environment [41]. BEL has two main inputs: sensory inputs (SI) and emotional signal (ES). It is worth mentioning that this method is an appealing strategy for addressing real-time applications with multi-objective properties due to its flexibility in defining both SI and ES [41]. Furthermore, BEL has a computational complexity in the order of $O(n)$ which is much more efficient than other existing learning algorithms [41, 44, 51].

Contributions

The contribution of this paper has three main components, specifically (i) a path planning and control co-design, (ii) a learning based approach that can effectively handle the uncertainties from unstable UAV system and complex environment, and (iii) a low computational learning technique that can be implemented in real-

time. To this end, we utilize the computational model of emotional learning in mammal's brain, i.e., BEL, for developing a novel path planning and intelligent control co-design for practical real-time implementation in UAVs in the presence of system uncertainties and dynamic environments. To the best of the authors' knowledge, this is the first time that BEL is implemented for accomplishing intelligent path planning and control co-design of UAVs. The learning capabilities introduced by the proposed approach to the path planning and coordination of UAVs, enhances the overall system robustness and performance, which is critical for different real-time applications e.g., emergence response, outer space exploration, etc. especially when dealing with challenges caused by dynamic and uncertain environments with multiple unknown moving obstacles. Moreover, the proposed method is feasible and very promising for implementing into diverse real-time applications due to its specific low computational complexity. Particularly, the developed BEL based design provides a path planner with multi-objective properties i.e., optimizing the path planning, improving the robustness and adaptivity for path planning which can deal with dynamic and uncertain environments, and having the capability to handle the impacts from multiple unknown moving obstacles. In addition, the proposed approach keeps the overall computational complexity of the system within a reasonable region. The main objective is then to design a practically applicable real-time path planner that is able of keeping the UAV performance as satisfactory as possible in terms of efficiently determining where the UAV should move from its starting and/or current position to a target location, while avoiding any collisions between the UAV and potential obstacles, and also satisfactory tracking a target. In order to demonstrate the effectiveness of the proposed approach, a set of numerical simulations are provided. A comparison between the proposed approach and other recent developed path planning strate-

gies is included, where it is possible to observe the path planning improvement produced by BEL-inspired technique.

The rest of the paper is organized as follows. Section 6.2 presents the problem formulation and relevant preliminaries about path planning and BEL technique. The proposed design is introduced in Section 6.3, which consists of a BEL based path planning and intelligent control co-design. Section 6.4 presents numerical simulation results. The conclusion of the paper and future directions of our work are provided in Section 6.5.

6.2 Problem Formulation and Preliminaries

Consider the real-time autonomous flight of a UAV in dynamic uncertain environments. The dynamics of the UAV are nonlinear and complex and, as expected from any real-time platform, the system dynamics are affected by uncertainties, noise, and disturbances. The main objective of this research consists on the design of a path planning and intelligent control co-design, which is effective for real-time autonomous flight of a UAV, even under the system uncertainties and dynamic environment.

The solution proposed in this paper is a novel path planning and intelligent control co-design, which provides the following benefits:

- optimizing the path planning.
- improving the robustness and adaptivity for path planning which allows dealing with dynamic and uncertain environments.

- capability to handle the impacts from unpredictable and unknown moving obstacles.
- suitable for real-time implementation due to its low computational complexity.
- ensuring the stability of the system during its online training phase.

The problem formulation and relevant preliminaries about path planning, UAV modeling, and BEL technique are provided next.

6.2.1 Path Planning in Continuous Spaces

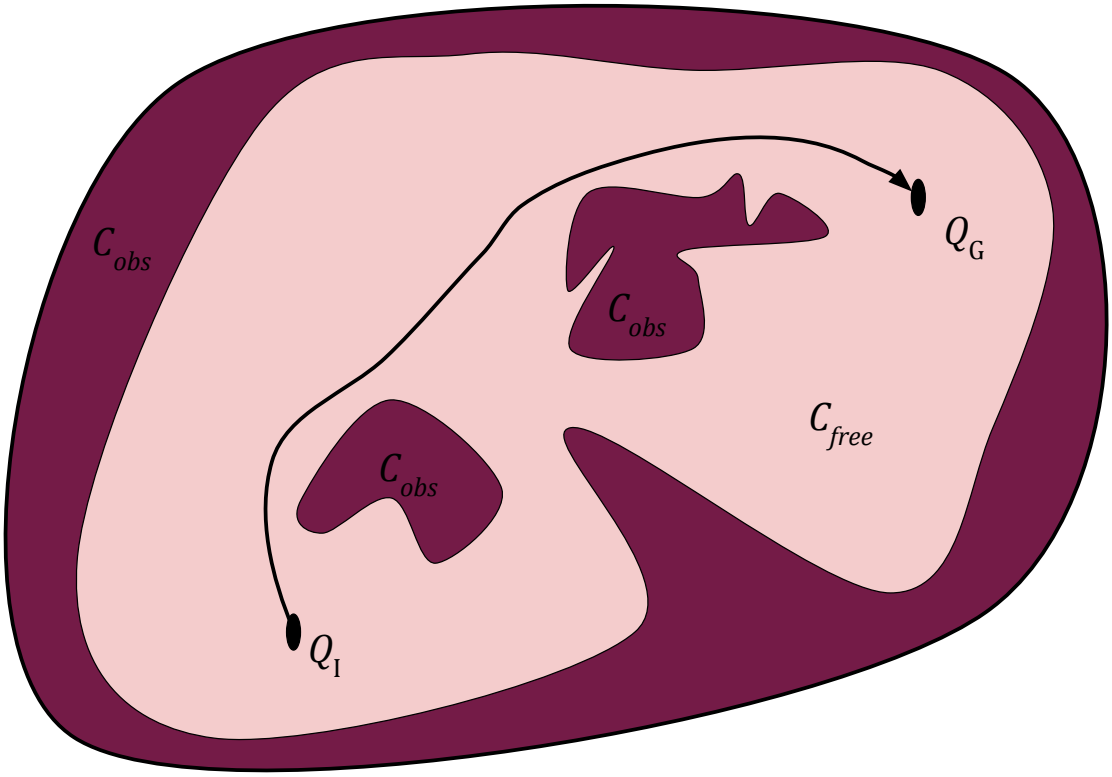

Figure 6.1: A path planning problem using configuration space (i.e., $\mathcal{C}$–space) ideas.

Consider a robot moving in an m–dimensional workspace $\mathcal{W} \subset \mathbb{R}^m$ and $(m = 2, 3)$ which contains a set of obstacles $\mathcal{O} \in \mathcal{W}$. Consider also that $\mathcal{Q} \in \mathcal{C}$ represents the configuration of robot $\mathcal{A}(\mathcal{Q})$, in which $\mathcal{Q} = (x, y, \psi)$ and $\mathcal{Q} = (x, y, z, \phi, \theta, \psi)$ for 2–dimensional and 3–dimensional workspace, respectively. Then, the obstacle region which is a closed set in $\mathcal{C}$, and the free space which is an open set, could be described as follows [122]:

$$\mathcal{C}_{obs} = \{\mathcal{Q} \in \mathcal{C} \mid \mathcal{A}(\mathcal{Q}) \cap \mathcal{O} \neq \varnothing\} \ , \ \mathcal{C}_{obs} \subseteq \mathcal{C}$$
$$\mathcal{C}_{free} = \mathcal{C} \backslash \mathcal{C}_{obs}$$

A typical path planning problem is shown in Figure 6.1. The main objective is to find a path from the start configuration ($\mathcal{Q}_I$) to the goal configuration ($\mathcal{Q}_G$) in $\mathcal{C}_{free}$. The entire map represents $\mathcal{C} = \mathcal{C}_{free} \cup \mathcal{C}_{obs}$. In other words, a complete path planning algorithm must be able to compute a continuous path (i.e., $\tau : [0, 1] \rightarrow \mathcal{C}_{free}$), in a way that $\tau(0) = \mathcal{Q}_I$ and $\tau(1) = \mathcal{Q}_G$. Also, if there is no solution for the path planning problem, it must be able to correctly report that such a path does not exist.

Various algorithms have been proposed for solving the path planning problem [122], where they could be categorized as one (or combinations) of the following methods: (i) sampling-based planning methods, (ii) combinatorial planning methods, (iii) optimal planning methods, and (iv) potential field planning methods. However, most of these methods exhibit issues like the problem of computational complexity, trapping in local minima, failing to find the solution, etc [122].

To overcome these deficiencies, we utilize the computational model of emotional learning in mammal's brain, i.e., BEL, for developing a novel path planning and intelligent control co-design for real-time implementation in UAVs in the pres-

ence of system uncertainties and dynamic environments.

6.2.2 UAV Dynamics Representation

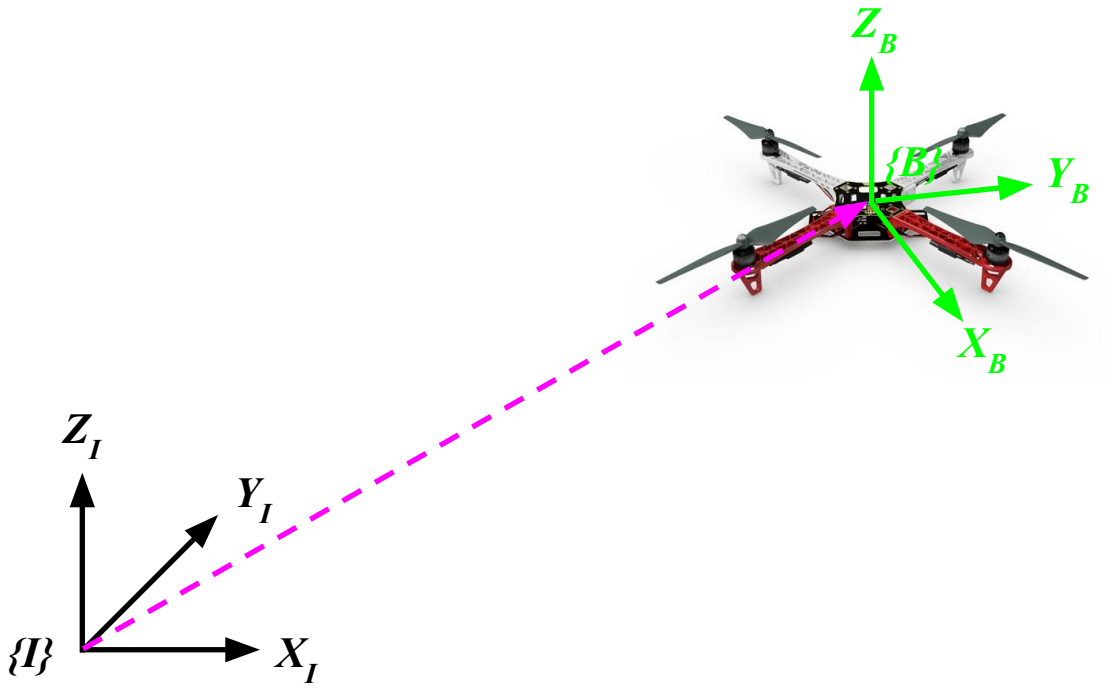

Figure 6.2: System Coordinates.

In order to analytically validating the proposed controller, the dynamics of the UAV are described by means of a hybrid system of coordinates, i.e., the H-frame [77]. This system is composed of a set of angular equations with respect to (w.r.t.) a body fixed frame (B-frame) and a set of nonlinear equations w.r.t. an inertial frame (I-frame), see Figure 6.2. The UAV nonlinear equations w.r.t. the H-frame are described as follows [77]:

$$
\begin{cases}
\ddot{X} = (\sin\psi\sin\phi + \cos\psi\sin\theta\cos\phi)\frac{U_1}{m} \\
\ddot{Y} = (-\cos\psi\sin\phi + \sin\psi\sin\theta\cos\phi)\frac{U_1}{m} \\
\ddot{Z} = -g + (\cos\theta\cos\phi)\frac{U_1}{m} \\
\dot{p} = \frac{I_{YY}-I_{ZZ}}{I_{XX}}qr - \frac{J_{TP}}{I_{XX}}q\mathbf{\Omega} + \frac{U_2}{I_{XX}} \\
\dot{q} = \frac{I_{ZZ}-I_{XX}}{I_{YY}}pr - \frac{J_{TP}}{I_{YY}}p\mathbf{\Omega} + \frac{U_3}{I_{YY}} \\
\dot{r} = \frac{I_{XX}-I_{YY}}{I_{ZZ}}pq + \frac{U_4}{I_{ZZ}}
\end{cases}
\tag{6.1}
$$

Additionally, the propellers' speed are related to the moments as follows:

Table 6.1: Definition of system parameters and variables

Parameters	Definitions
θ	Pitch angle
ϕ	Roll angle
ψ	Yaw angle
U_1	collective throttle
U_2	Roll moment
U_3	Pitch moment
U_4	Yaw moment
I_{XX}	body moment of inertia around the X axis
I_{YY}	body moment of inertia around the Y axis
I_{ZZ}	body moment of inertia around the Z axis
g	acceleration due to gravity
m	UAV mass
J_{TP}	total rotational moment of inertia around the propeller axis
$\mathbf{\Omega}_1$	front-left propeller speed
$\mathbf{\Omega}_2$	front-right propeller speed
$\mathbf{\Omega}_3$	rear-right propeller speed
$\mathbf{\Omega}_4$	rear-left propeller speed
b_q	thrust factor
d_q	drag factor
l	distance between center of the quadrotor and center of the propeller
K_E	electric motor constant
K_M	mechanic motor constant
R	motor resistance
ν	motor voltage
N	gearbox reduction ratio
η	conversion efficiency of the gearbox

$$
\begin{cases}
U_1 = b_q(\mathbf{\Omega}_1^2 + \mathbf{\Omega}_2^2 + \mathbf{\Omega}_3^2 + \mathbf{\Omega}_4^2) \\
U_2 = b_q l(-\mathbf{\Omega}_2^2 - \mathbf{\Omega}_3^2 + \mathbf{\Omega}_1^2 + \mathbf{\Omega}_4^2) \\
U_3 = b_q l(-\mathbf{\Omega}_1^2 - \mathbf{\Omega}_2^2 + \mathbf{\Omega}_3^2 + \mathbf{\Omega}_4^2) \\
U_4 = d_q(-\mathbf{\Omega}_1^2 + \mathbf{\Omega}_2^2 - \mathbf{\Omega}_3^2 + \mathbf{\Omega}_4^2) \\
\mathbf{\Omega} = -\mathbf{\Omega}_1 + \mathbf{\Omega}_2 - \mathbf{\Omega}_3 + \mathbf{\Omega}_4
\end{cases}
\tag{6.2}
$$

Ultimately, the nonlinearities of the quad rotorcraft motors is described as follows:

$$J_{TP}\dot{\mathbf{\Omega}} = -\frac{K_E K_M}{R}\eta N^2 \mathbf{\Omega} - d_q \mathbf{\Omega}^2 + \frac{K_M}{R}\eta N \nu \tag{6.3}$$

The parameters of Equations (6.1), (6.2), and (6.3) are given in Table 6.1.

6.2.3 Brain Emotional Learning

Brain Emotional Learning (BEL) is a neurobiologically-motivated intelligent methodologies, which is based on the computational model of emotional learning observed in the mammalian limbic system [40]. This model, depicted in Figure 6.3, has two main parts: *Amygdala*, and *Orbitofrontal Cortex*. Amygdala is responsible for immediate learning, while Orbitofrontal Cortex is responsible for inhibition of any inappropriate learning happening in the Amygdala. *Sensory Inputs* (SI) and *Emotional Signal* (ES) are two main inputs to the BEL model.

The BEL model output (MO) can be defined as

$$MO = \sum_l A_l - \sum_l OC_l \tag{6.4}$$

which is calculated by the difference between the Amygdala outputs (A_l) and the Orbitofrontal Cortex outputs (OC_l). Here, l is the number of sensory inputs.

The Orbitofrontal Cortex and the Amygdala outputs are calculated by the summation of all their corresponding nodes, where the output of each node is de-

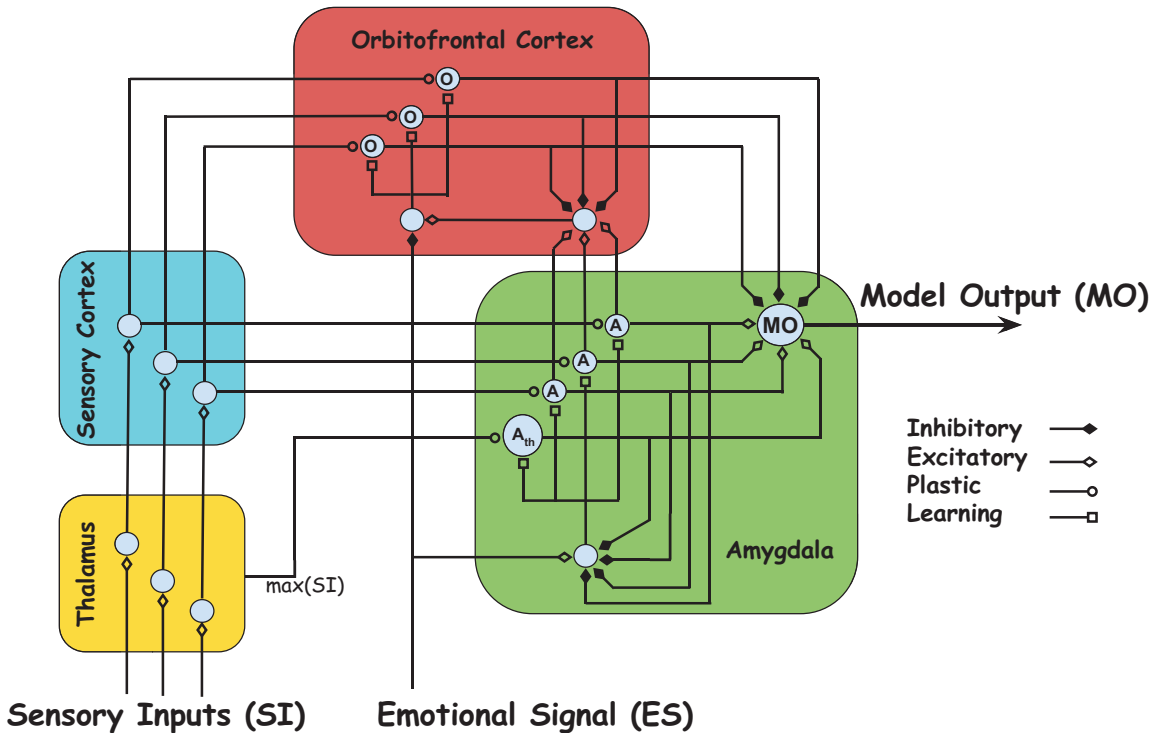

Figure 6.3: Computational model of emotional learning.

scribed as:

$$A_l = V_l SI_l \tag{6.5}$$

$$OC_l = W_l SI_l \tag{6.6}$$

where SI_l is the l^{th} sensory input, V_l is the weight of the Amygdala, and W_l is the weight of the Orbitofrontal Cortex. The following equations are employed for updating V_l and W_l, respectively:

$$\Delta V_l = K_v SI_l \max\left(0, ES - \sum_l A_l\right) \tag{6.7}$$

$$\Delta W_l = K_w SI_l \left(MO - ES\right) \tag{6.8}$$

where, K_w and K_v are the learning rates.

The maximum of all sensory inputs is another input considered in the model. This signal, which is directly sent from the Thalamus to the Amygdala, is defined as:

$$A_{th} = V_{th} \max(SI_l) \tag{6.9}$$

with V_{th} being the weight, and with a corresponding update law similar to the one shown in Equation (6.7).

Several techniques have been adopted for tuning the BEL parameters. For instance, genetic algorithm (GA) is adopted for optimally tuning BEL parameters in [130] while a particle swarm optimization-based approach is implemented in [131]. The authors in [47] adopted the clonal selection algorithm to obtain BEL parameters, where it has been successfully applied for controlling a single-link flexible joint manipulator. Moreover, a fuzzy tuning of BEL parameters has been proposed in [132], and successfully applied for controlling a chaotic system and an inverted double pendulum system. Trial and error tuning has also shown to be appropriate, since [41] and [51] relied on this method. In this paper, to significantly reduce the computational complexity, a heuristic approach is utilized for tuning the BEL parameters.

6.2.4 Objectives

Based on the quad rotorcraft model described in Section 6.2.2, and by leveraging the brain emotional learning introduced in Section 6.2.3, the objective is to design a path planning and intelligent control co-design for UAVs. Specifically, the proposed method is designed for practical real-time implementation in UAVs in presence of system uncertainties and dynamic environments.

6.3 Brain Emotional Learning-Based Path Planning and Intelligent Control Co-Design

6.3.1 System Design

The primary idea behind the emotional learning based path planning and/or decision-making and intelligent control co-design in this work is to produce the action (i.e., output) that regulates the emotional signal i.e., the one that maximizes the emotional reward or minimizes the emotional stress, while different sets of sensory inputs are received by the system [41]. The current situation of the system is represented by the received sensory inputs while the emotional signals represent how satisfactory the performance of the system is. In other words, the emotional signal represents the condition of the system by considering the particular objective.

In this paper, we propose a novel path planning and intelligent control co-design which is based on the computational model of emotional learning in the mammal's brain. The BEL-inspired path planner aims at a practical real-time implementation in UAVs in presence of system uncertainties and dynamic environments.

It is worth mentioning that the proposed path planner will intelligently learn in real-time the best path for the UAV, and will provide this information to the control layer, in such a way it can be used as a control objective. After receiving the latest learnt path, the BEL-based control algorithm will force the UAV to achieve these objectives i.e., track the desired path, save the energy etc. in an intelligent manner.

Path Planner

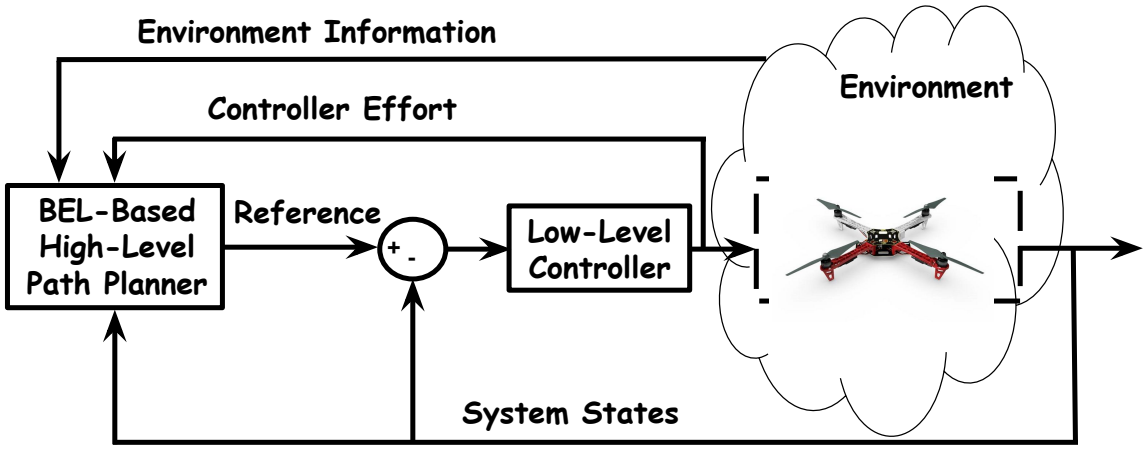

Figure 6.4: BEL-based path planning architecture.

The design of a BEL-based path planner, is desirable since it could be practically accomplished without increasing the complexity of the overall system. The BEL-based path planning architecture implemented in this work is shown in Figure 6.4. This figure demonstrates a closed loop configuration which consists of the following blocks: (i) BEL-based high-level path planner block, (ii) low-level controller block, (iii) environment block, and finally (iv) a block for the UAV. This architecture implicitly demonstrates the overall emotional learning based path planning concept, which consists of the action selection mechanism, the critic, and the learning algorithm.

Intelligent Control Co-Design

In general, intelligent control techniques can be utilized for solving different control problems via direct or indirect approaches. In the direct mode, the intelligent method is utilized as a controller block, while it is employed for obtaining the controller's parameters in the indirect mode. In this paper, we propose an intelligent

control co-design based on the computational model of emotional learning in the mammal's brain, which is utilized as a low-level controller for path planning and intelligent control co-design of a UAV.

Aiming at designing an intelligent control co-design well suited for implementation in real-time systems, the proposed methodology aims at improving: (i) target tracking, (ii) disturbance rejection, and (iii) model uncertainty handling. To fulfill these objectives, each of the control inputs $\{U_1,U_2,U_3,U_4\}$, will be designed as [48]:

$$\begin{aligned} SI_i^{UAV} &= K_{i1}e + K_{i2}\int e \times dt + K_{i3}\frac{de}{dt} \\ ES_i^{UAV} &= K_{i4}e + K_{i5}\int e \times dt + K_{i6}\frac{de}{dt} \\ U_i &= \sum_i V_i^{UAV} \times SI_i^{UAV} \\ &\quad - \sum_i W_i^{UAV} \times SI_i^{UAV} \end{aligned} \tag{6.10}$$

Here, $i = 1, ..., 4$ makes reference to each control input. K_{i1}, K_{i2}, K_{i3}, K_{i4}, K_{i5}, and K_{i6} are positive gains. By assigning different values to these gains, the ES will change its influence on the system behavior. In this work, different values are used for each one of control inputs of the system, i.e., for U_i, $i = 1, ..., 4$.

6.3.2 Emotional Signal and Sensory Input Development

Fundamentally, BEL is an action selection technique, in which an action is produced based on the sensory input (SI) and the emotional signal (ES). The general forms of SI and ES are given as follows:

$$SI = G\left(SS, EI, U\right) \tag{6.11}$$

$$ES = F\left(SS, EI, U\right) \tag{6.12}$$

where SS is the system states vector, EI is the environment information vector, and U is the control effort vector.

The path planning objectives (e.g., optimizing the path planning and robustness and adaptivity) can be decided implicitly by choosing the adequate ES. For example, it is possible to choose the ES for achieving a better robustness and adaptivity performance, for achieving the optimal path, and/or for the energy expense minimization, among others.

Since our goal is to design a path planning and intelligent control co-design for UAVs, the proposed biologically inspired co-design technique will focus on: (i) optimizing the path planning, (ii) improving the robustness and adaptivity for path planning which can deal with dynamic and uncertain environments, and (iii) having the capability to handle the impacts from unpredictable and unknown moving obstacles.

To accomplish these objectives, the SI and ES, will be designed as:

$$\begin{aligned} SI = {} & K_{SI}^{1}U_1 + K_{SI}^{2}U_2 + K_{SI}^{3}U_3 \\ & + K_{SI}^{4}U_4 + K_{SI}^{5}U_{PF} + K_{SI}^{6}\Upsilon\left(EI\right) \end{aligned} \tag{6.13}$$

$$\begin{aligned} ES = {} & K_{ES}^{1}U_1 + K_{ES}^{2}U_2 + K_{ES}^{3}U_3 \\ & + K_{ES}^{4}U_4 + K_{ES}^{5}U_{PF} + K_{ES}^{6}\Upsilon\left(EI\right) \end{aligned} \tag{6.14}$$

where $U_i, i = \{1, ...4\}$ are the UAV controllers effort, K_{SI}^{l}, $K_{ES}^{l}, l = \{1, ...6\}$, are positive gains, $\Upsilon\left(EI\right)$ is a function representing the environment information and U_{PF} is a potential field function. The ES will change its impact on the system be-

havior by assigning different values to these positive gains. In this work, different gains are assigned for each one of the ES and SI.

In this work, the ES is designed in such a way that the increase in reference tracking error will generate a *negative emotion* in the system, which is then taken as an evidence for the unsatisfactory performance of the system. Therefore, the proposed planner will behave in such a way that it will always minimize the negative emotion, ultimately leading to the satisfactory performance of the system.

6.3.3 Learning-based Path Planner

In path planning of UAVs, multiple performance considerations have to be taken into account all at the same time, therefore, it is a very interesting case for using biologically-inspired learning-based multi-objective methodologies like BEL. Designing a path planner suitable for real- time implementation, as well as an intelligent control co-design for UAVs, encourages us to take advantage of the computational model of emotional learning in the mammals' limbic system, i.e., BEL.

From equations (6.13)-(6.14), the BEL-based path planner for UAVs is defined as

$$\begin{aligned} u^{BEL} &= V \times SI - W \times SI \\ &= V \times (K_{SI}^1 U_1 + K_{SI}^2 U_2 + K_{SI}^3 U_3 \\ &+ K_{SI}^4 U_4 + K_{SI}^5 U_{PF} + K_{SI}^6 \Upsilon(EI)) \\ &- W \times (K_{SI}^1 U_1 + K_{SI}^2 U_2 + K_{SI}^3 U_3 \\ &+ K_{SI}^4 U_4 + K_{SI}^5 U_{PF} + K_{SI}^6 \Upsilon(EI)) \end{aligned} \tag{6.15}$$

Substituting the Emotional Signal with equation (6.14) the BEL model output of the path planner for UAVs could be obtained as follows:

$$\begin{aligned} MO &= ES \\ &= K_{ES}^{1}U_1 + K_{ES}^{2}U_2 + K_{ES}^{3}U_3 \\ &+ K_{ES}^{4}U_4 + K_{ES}^{5}U_{PF} + K_{ES}^{6}\Upsilon\left(EI\right) \end{aligned} \tag{6.16}$$

which clearly satisfies our goal of path planning for UAVs.

The overall path planner proposed in this paper is summarized as pseudo-code in **Algorithm** 5.

Algorithm 5 : The BEL-based methodology for path planning and intelligent control co-design of UAVs.

Initialization:
Set $W^{UAV} = 0$, $V_{th}^{UAV} = 0$, and $V^{UAV} = 0$.
Set $W = 0$, $V_{th} = 0$, and $V = 0$.
Define $ES =$ Objective function.
for each iteration $t = t_s$ **do**
  for each control inputs U_i, $i = 1, ..., 4$ **do**
    Compute
    $U_i = V_i^{UAV} \times SI_i^{UAV} - W_i^{UAV} \times SI_i^{UAV}$
    Compute SI_i^{UAV}
    Compute ES_i^{UAV}
    Update V_i^{UAV}
    Update $V_{th,i}^{UAV}$
    Update W_i^{UAV}
  end for
  Compute $SI = K_{SI}^{1}U_1 + K_{SI}^{2}U_2 + K_{SI}^{3}U_3 + K_{SI}^{4}U_4 + K_{SI}^{5}U_{PF} + K_{SI}^{6}\Upsilon\left(EI\right)$
  Compute $ES = K_{ES}^{1}U_1 + K_{ES}^{2}U_2 + K_{ES}^{3}U_3 + K_{ES}^{4}U_4 + K_{ES}^{5}U_{PF} + K_{ES}^{6}\Upsilon\left(EI\right)$
  Compute $A = V \times SI$
  Compute $A_{th} = V_{th} \times \max\left(SI\right)$
  Compute $OC = W \times SI$
  Compute $BEL = A - OC$
  Update V
  Update V_{th}
  Update W
end for

6.3.4 Stability Analysis

Theorem 8 presents the convergence of the weights of the Amygdala (V_l) and the Orbitofrontal Cortex (W_l). Next, Remark 10 explains how the proposed method accomplishes a path planner for UAVs.

Theorem 8. *Given the BEL as shown in Equations* (6.13)–(6.16)*, there exists positive BEL tuning parameters, K_v, K_w satisfying*

I. $\left|\left[1 - K_v \left(SI_l\right)^2\right]\right| < 1$

II. $\left|\left[1 - K_w \left(SI_l\right)^2\right]\right| < 1$

such that the BEL estimated weights of the Amygdala (V_l) and the Orbitofrontal Cortex (W_l) converge to the desired targets asymptotically.

Proof. See **Appendix A** □

Remark 10. *Based on the BEL theory [41] and making use of Equation* (6.15)*, the path planner for UAVs can be obtained while the estimated weights of the Amygdala (V_l) and the Orbitofrontal Cortex (W_l) are converging to desired targets. According to Theorem 8, the estimated weights converge to the desired targets asymptotically. Therefore, the designed BEL input U shown in Equation* (6.15) *asymptotically accomplishes an optimal path planner for UAVs.*

6.4 Simulation Results

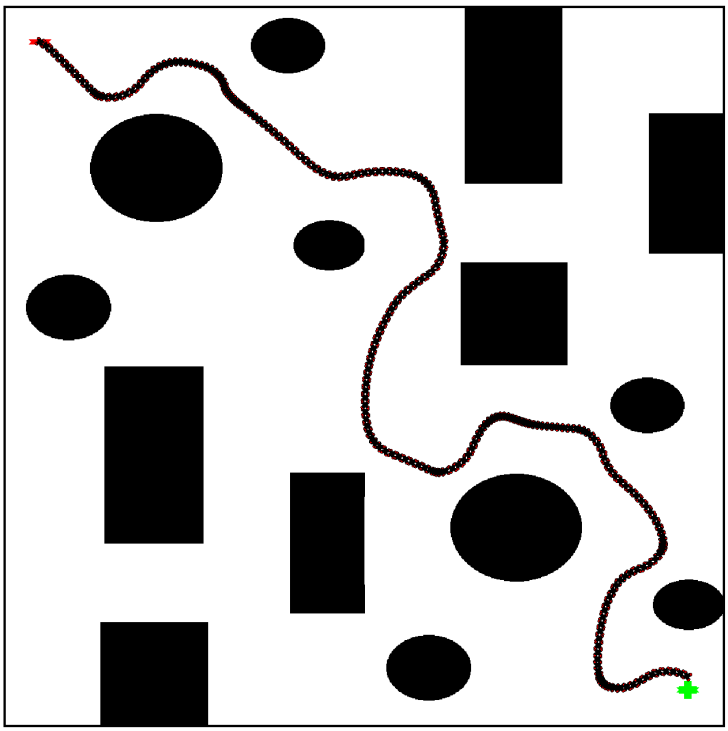

Figure 6.5: Path generated by the BEL-based path planner for workspace with multiple obstacles with different shape and size.

This section presents computer-based simulation results showing the performance of the proposed path planning and intelligent control co-design for unmanned aerial vehicles. Figure 6.5 shows the results of applying the proposed method. The initial position is in red and the goal position is plotted in green.

Table 6.2: Comparison of BEL-based path planning and RRT

Algorithms	Processing Time	Path Length
BEL-based	$6.419738E+01$	$1.874101E+03$
RRT	$8.182490E+01$	$1.567674E+03$

For comparison purposes, a similar experiment was performed, but using the RRT algorithm instead of the proposed algorithm. Figure 6.6 shows the results

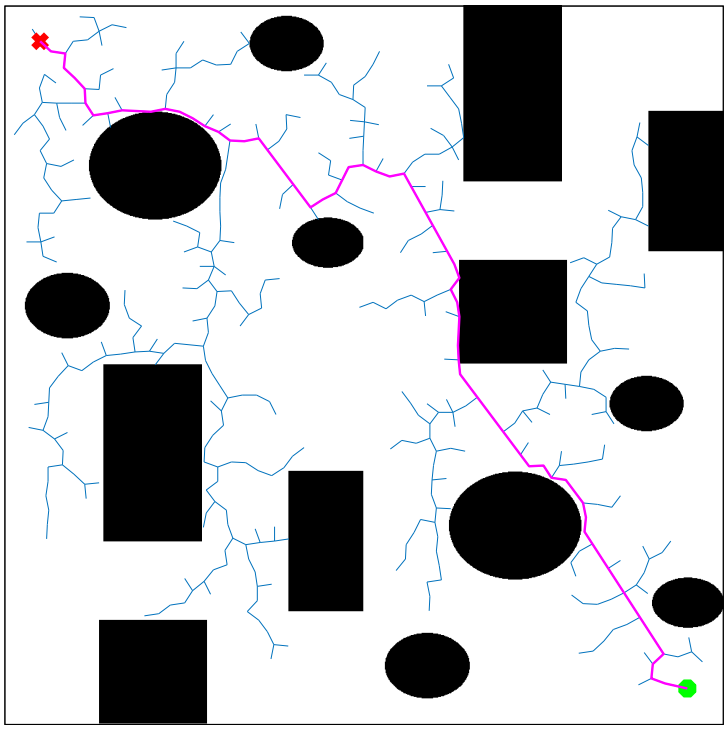

Figure 6.6: Path generated by the RRT-based path planner for workspace with multiple obstacles with different shape and size.

of applying the RRT method in the same workspace. Table 6.2 summarizes some relevant outcomes for each one of the two methods.

Furthermore, a similar experiment was performed, in the same workspace, but using a conventional PID controller instead of the intelligent control co-design. The results of applying the PID method are shown in Figure 6.7.

Experimental Testbed

The platform implemented for future validation of the proposed algorithm is available at the Unmanned Systems Laboratory (USL) from the University of Nevada,

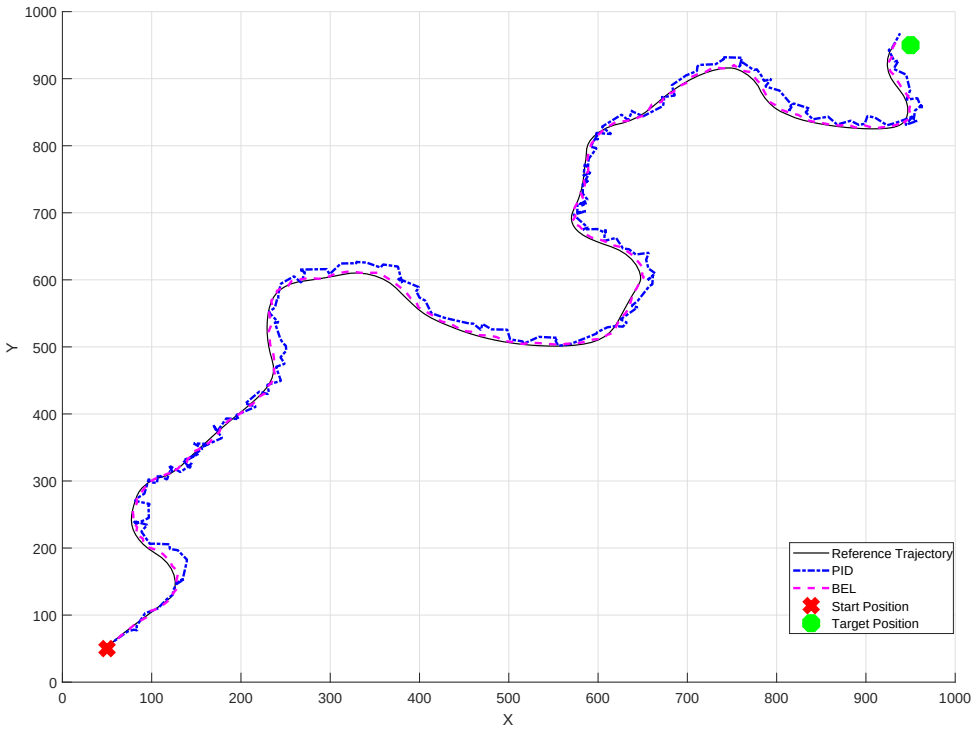

Figure 6.7: Trajectory tracking by the BEL-based control co-design (in Magenta Color) and PID-based co-design (in Blue Color) for workspace with multiple obstacles with different shape and size.

Reno. The Base Station of this testbed runs Ubuntu 14.04 OS, the Robot Operating System (ROS) environment, and Matlab. The UAV platform corresponds to a Bebop drone manufactured by Parrot.

The 3-dimensional position of the UAV is obtained by means of a Motion Capture System (MCS) manufactured by OptiTrack. The information provided by the MCS is reported to the OptiTrack Interface PC by means of a Gigabyte Ethernet connection. Next, this information is sent to the Base Station PC by means of an Ethernet connection. The Base Station computer uses this information to execute the BEL algorithm and to calculate the path planing and intelligent control co-design, which are sent to the Bebop platform by means of a WiFi link. Figure 6.8 shows the experimental testbed which will be used for future evaluation of the

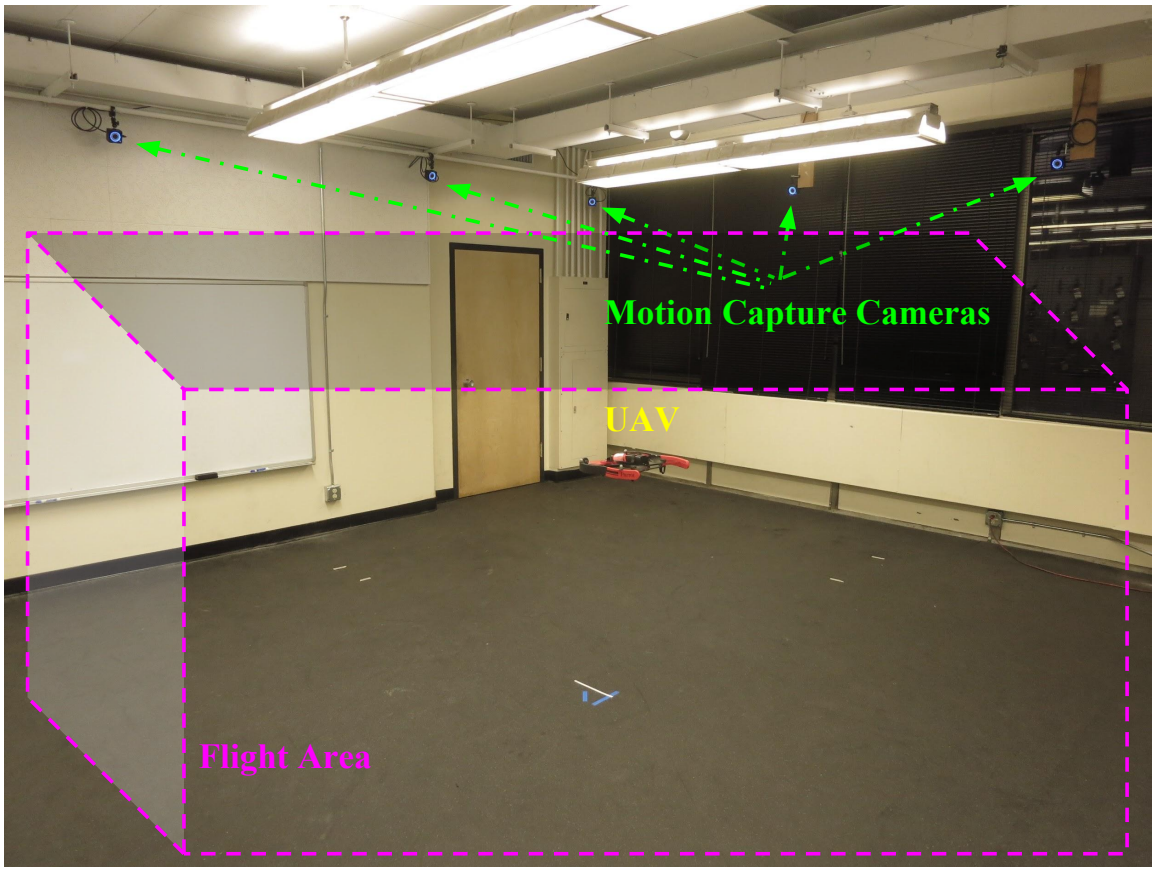

Figure 6.8: The experimental testbed for future evaluation of the proposed algorithm on a Bebop drone manufactured by Parrot.

proposed algorithm on a Bebop drone.

6.5 Conclusions and Future Works

In this paper, a joint intelligent path planning and control co-design for UAVs in the presence of system uncertainties and dynamic environment has been investigated. The proposed approach, which makes use of a biologically-inspired methodology know as BEL was implemented for the first time in this application. The novel co-design enhanced the UAV for dealing with the challenges caused by system uncertainties and dynamic environments. Furthermore, the proposed method is

very promising for implementation in real-time applications due to its specific low computational complexity. Numerical results of the BEL-based path planning and intelligent control co-design for UAVs demonstrate the effectiveness of the proposed approach.

6.6 Appendix A

6.6.1 Non-adapting Phase

Our goal is to investigate the output of the system in non-adapting phase (i.e., when the system completes its learning process). In addition we make an assumption that the $max(\cdot)$ function in equation (6.7) could be neglected. By substituting equations (6.5) and (6.6) in equation (6.4) the output of the model could be defined as follows:

$$\begin{aligned} MO =& \sum_{l} V_l SI_l - \sum_{l} W_l SI_l \\ =& \sum_{l} \left(V_l - W_l\right) SI_l \end{aligned} \tag{6.17}$$

$$\begin{aligned} \Delta V_l =& K_v SI_l \left(ES - \sum_{l} A_l \right) \\ =& K_v SI_l \left(ES - \sum_{l} V_l SI_l \right) \end{aligned} \tag{6.18}$$

$$\begin{aligned} \Delta W_l =& K_w SI_l \left(MO - ES\right) \\ =& K_w SI_l \left(\sum_{l} \left(V_l - W_l\right) SI_l - ES \right) \end{aligned} \tag{6.19}$$

When the learning process is completed i.e., after system completes its learning process the variations of the weights of Amygdala (ΔV_l) and Orbitofrontal Cortex (ΔW_l) will be equal to zero (i.e., $\Delta V_l = \Delta W_l = 0$). With the assumption of $SI_l \neq 0$ the following holds:

$$\begin{aligned} K_v SI_l \left(ES - \sum_l V_l SI_l \right) &= 0 \\ \Rightarrow \sum_l V_l SI_l &= ES \end{aligned} \tag{6.20}$$

$$\begin{aligned} K_w SI_l \left(\sum_l \left(V_l - W_l \right) SI_l - ES \right) &= 0 \\ \Rightarrow \sum_l \left(V_l - W_l \right) SI_l &= ES \\ \Rightarrow \sum_l W_l SI_l &= 0 \end{aligned} \tag{6.21}$$

By substituting equations (6.20) and (6.21) in equation (6.17) the model output in non-adapting phase will be as follows:

$$MO = \sum_l \left(V_l - W_l \right) SI_l = \sum_l V_l SI_l = ES \tag{6.22}$$

6.6.2 Main Proof

Considering the results obtained in Subsection 6.6.1 the following should be achieved:

$$MO_l \rightarrow ES_l \tag{6.23}$$

Let V_l^* be the weight of Amygdala for each control input l when the system has been learned and let $\hat{ES_l}$ be the Emotional Signal for each control input l during the adaptation phase. The following equations hold:

$$ES_l = V_l^* SI_l \text{ and } \hat{ES_l} = V_l SI_l \tag{6.24}$$

$$\Delta V_l(k) = K_v SI_l \max\left(0, ES_l - \hat{ES_l}\right) \tag{6.25}$$

We will investigate the results of the following two cases:

Case I. $ES_l - \hat{ES_l} \geq 0$

Case II. $ES_l - \hat{ES_l} < 0$

Considering the case I., the proof can be achieved as follows:

$$\begin{aligned}
\Delta V_l(k) =& K_v SI_l \max\left(0, ES_l - \hat{ES_l}\right) \\
=& K_v SI_l \left(ES_l - \hat{ES_l}\right) \\
=& K_v SI_l \left(V_l^* SI_l - V_l SI_l\right) \\
=& K_v SI_l \left(V_l^* - V_l\right) SI_l \\
=& K_v SI_l \tilde{V_l} SI_l \\
=& K_v \tilde{V_l} \left(SI_l\right)^2
\end{aligned} \tag{6.26}$$

where $\tilde{V_l} = V_l^* - V_l$.

$$\begin{aligned}
V_l\left(k+1\right) =& V_l\left(k\right) + \Delta V_l(k) \\
\tilde{V_l}\left(k+1\right) =& V^* - V_l\left(k\right) - \Delta V_l(k) \\
=& \tilde{V_l}\left(k\right) - K_v \tilde{V_l}\left(SI_l\right)^2 \\
=& \left[1 - K_v\left(SI_l\right)^2\right] \tilde{V_l}\left(k\right)
\end{aligned} \tag{6.27}$$

Considering the case II., it is obvious that when $ES_l - \hat{ES_l} < 0$ the max function in equation (6.25) will force the adaptation in Amygdala to stop and the following holds:

$$\begin{aligned}
\Delta V_l(k) =& 0 \\
V_l\left(k+1\right) =& V_l\left(k\right) \\
\tilde{V_l}\left(k+1\right) =& \tilde{V_l}\left(k\right)
\end{aligned} \tag{6.28}$$

The proof can be achieved as follows:

$$\begin{aligned}
\Delta W_l(k) =& K_w SI_l \left(MO_l - ES_l\right) \\
=& K_w SI_l \left(V_l SI_l - W_l SI_l - V_l^* SI_l\right) \\
=& K_w SI_l \left(-\left(V_l^* - V_l\right) SI_l - W_l SI_l\right) \\
=& K_w SI_l \left(\left(-\tilde{V}_l - W_l\right) SI_l\right) \\
=& - K_w \tilde{W}_l \left(SI_l\right)^2
\end{aligned} \tag{6.29}$$

where $\tilde{V}_l = V_l^* - V_l$ and $\tilde{W}_l = \tilde{V}_l + W_l$.

$$\begin{aligned}
W_l\left(k+1\right) =& W_l\left(k\right) + \Delta W_l(k) \\
\tilde{W}_l\left(k+1\right) =& \tilde{V}_l\left(k+1\right) + W_l\left(k+1\right) \\
=& \tilde{V}_l\left(k\right) + W_l\left(k\right) + \Delta W_l(k) \\
=& \tilde{W}_l\left(k\right) - K_w \tilde{W}_l \left(SI_l\right)^2 \\
=& \left[1 - K_w \left(SI_l\right)^2\right] \tilde{W}_l\left(k\right)
\end{aligned} \tag{6.30}$$

Chapter 7

Conclusions and Future Directions

In this dissertation a biologically-inspired intelligent controller based on a novel architecture of emotional learning in mammal's brain was proposed for distributed control of MAS. The methodology, which is a biologically-inspired reinforcement learning, was designed and implemented in the real-time coordination of multiple UAS platforms operating in presence of system uncertainties and dynamic environment.

7.1 Conclusions

7.1.1 A Neurobiologically-inspired Intelligent Trajectory Tracking Control for Unmanned Aircraft Systems with Uncertain System Dynamics and Disturbance

This paper addressed the problem of stabilizing the full 6 DoF of a quad rotorcraft UAS and also intelligent tracking control of a UAS subjected to unknown system dynamics and external disturbances. The low-computational model-free BELBIC, a neurobiologically-motivated intelligent controller, was adopted in order to design and experimentally validate a novel UAS control methodology. The numerical and experimental flight results, which considered uncertainty and disturbances, demonstrated the effectiveness, applicability, and superior performance of the BELBIC-inspired controller, when compared to conventional model-based control methods. In addition, the convergence analysis of the proposed approach has been studied.

7.1.2 A Biologically-Inspired Reinforcement Learning based Intelligent Distributed Flocking Control for Multi-Agent Systems in Presence of Uncertain System and Dynamic Environment

In this paper a biologically-inspired intelligent controller based on a novel architecture of emotional learning in mammal's brain was proposed for flocking control of MAS. The methodology, which is given the name BELBIC-based flocking, was designed and implemented in the real-time coordination of multiple UAS platforms operating in presence of system uncertainties and dynamic environment. In addition, the convergence analysis of the proposed approach has been studied. Numerical and experimental results of the BELBIC-inspired flocking demonstrated the effectiveness of the proposed approach, as well as its applicability to real-time systems.

7.1.3 A Biologically-Inspired Distributed Fault Tolerant Flocking Control for Multi-Agent System in Presence of Uncertain Dynamics and Unknown Disturbance

A neurobiologically-motivated intelligent distributed resilient controller based on a computational model of emotional learning in the mammalian limbic system was proposed for flocking control of MAS in presence of system uncertainties and unknown disturbances. The methodology, called R-BELBIC-inspired flocking, em-

beds a resilience mechanism with multi-objective properties into the flocking control strategy in a distributed manner. The results from both computer-aid simulation and experimental test demonstrate the effectiveness of proposed R-BELBIC based resilient distributed flocking control, as well as its applicability for real-time systems.

7.1.4 A Game Theoretic Based Biologically-Inspired Distributed Intelligent Flocking Control for Multi-UAV Systems with Network Imperfections

A game theoretic based biologically-inspired distributed intelligent control methodology is proposed to overcome the challenges of network-induced delay in flocking of networked multi-unmanned aerial vehicles. The methodology is adopted based on the emotional learning phenomenon in the mammalian limbic system. The learning capability and low computational complexity of the proposed technique makes it a promising tool for implementing in real-time networked multi-unmanned aerial vehicles flocking considering the influence of network-induced delay. Computer-based numerical results of the implementation of the proposed methodology demonstrate the effectiveness of this algorithm for distributed intelligent flocking control of networked multi-UAV.

7.1.5 Brain Emotional Learning-Based Path Planning and Intelligent Control Co-Design for Unmanned Aerial Vehicle in Presence of System Uncertainties and Dynamic Environment

In this paper, a joint intelligent path planning and control co-design for UAVs in the presence of system uncertainties and dynamic environment has been investigated. The proposed approach, which makes use of a biologically-inspired methodology know as BEL was implemented for the first time in this application. The novel co-design enhanced the UAV for dealing with the challenges caused by system uncertainties and dynamic environments. Furthermore, the proposed method is very promising for implementation in real-time applications due to its specific low computational complexity. Numerical results of the BEL-based path planning and intelligent control co-design for UAVs demonstrate the effectiveness of the proposed approach.

7.2 Future Directions

Several future work items are inline. In Chapter 2 the implementation of a BELBIC-inspired intelligent control strategy for addressing the task of UAS-based autonomous transportation of loads with uncertain characteristics will be considered for future work. In Chapters 3 and 4 one can extend our results also to handle other resources of faults such as network imperfections, actuator faults, etc. Also, in this dissertation, we assumed that all the parts using the same clock and has the time stamp for maintaining the synchronization. Since both the Amygdala and the Orbitofrontal

Cortex are located at the same layer in the BELBIC model, they could be considered to work with the same clock. Since the synchronization is a critical issue, and put all the parts in the same clock might not be the best solution. We will consider the synchronization issue in the future work. In Chapter 5 we took advantage of the game theory for designing the ES where every agent tried to minimize the total flocking error over a finite time interval, while at the same time minimizing its control effort. This work can be extended by designing a zero-sum game. Finally, in Chapter 6 a joint intelligent path planning and control co-design for UAVs in the presence of system uncertainties and dynamic environment has been proposed. Future works will consider improving the performance of the proposed method as well as providing the experimental results. Furthermore, extending the proposed method for applying to MAS will be under consideration.

Bibliography

[1] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Transactions on automatic control*, vol. 51, no. 3, pp. 401–420, 2006.

[2] H. M. La, R. Lim, and W. Sheng, "Multirobot cooperative learning for predator avoidance," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 1, pp. 52–63, 2015.

[3] G. Ortega, F. Mu, E. E. Quesada, L. R. Garcia, P. Ordaz, *et al.*, "Implementation of leader-follower linear consensus algorithm for coordination of multiple aircrafts," in *2015 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED-UAS)*, pp. 25–32, IEEE, 2015.

[4] H. Li, X. Liao, T. Huang, W. Zhu, and Y. Liu, "Second-order global consensus in multiagent networks with random directional link failure," *IEEE transactions on neural networks and learning systems*, vol. 26, no. 3, pp. 565–575, 2015.

[5] Z.-H. Guan, G.-S. Han, J. Li, D.-X. He, and G. Feng, "Impulsive multiconsensus of second-order multiagent networks using sampled position data," *IEEE transactions on neural networks and learning systems*, vol. 26, no. 11, pp. 2678–2688, 2015.

[6] H. Xu and L. R. G. Carrillo, "Distributed near optimal flocking control for multiple unmanned aircraft systems," in *Unmanned Aircraft Systems (ICUAS), 2015 International Conference on*, pp. 879–885, IEEE, 2015.

[7] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," *ACM SIGGRAPH computer graphics*, vol. 21, no. 4, pp. 25–34, 1987.

[8] X. Liu, J. Lam, W. Yu, and G. Chen, "Finite-time consensus of multiagent systems with a switching protocol," *IEEE transactions on neural networks and learning systems*, vol. 27, no. 4, pp. 853–862, 2016.

[9] Z. Peng, D. Wang, H. Zhang, and G. Sun, "Distributed neural network control for adaptive synchronization of uncertain dynamical multiagent systems," *IEEE transactions on neural networks and learning systems*, vol. 25, no. 8, pp. 1508–1519, 2014.

[10] C. P. Chen, G.-X. Wen, Y.-J. Liu, and F.-Y. Wang, "Adaptive consensus control for a class of nonlinear multiagent time-delay systems using neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 6, pp. 1217–1226, 2014.

[11] M. Jafari, "On the cooperative control and obstacle avoidance of multivehicle systems," Master's thesis, University of Nevada, Reno, 2015.

[12] W. Chen, S. Hua, and H. Zhang, "Consensus-based distributed cooperative learning from closed-loop neural control systems," *IEEE transactions on neural networks and learning systems*, vol. 26, no. 2, pp. 331–345, 2015.

[13] T. Nguyen, H. M. La, T. D. Le, and M. Jafari, "Formation control and obstacle avoidance of multiple rectangular agents with limited communication ranges," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 4, pp. 680–691, 2017.

[14] T. Nguyen, H. M. La, and M. Jafari, "On the formation control of a multivehicle system,"

[15] O. O'Loan and M. Evans, "Alternating steady state in one-dimensional flocking," *Journal of Physics A: Mathematical and General*, vol. 32, no. 8, pp. L99–L105, 1999.

[16] S. Li, X. Liu, W. Tang, and J. Zhang, "Flocking of multi-agents following a leader with adaptive protocol in a noisy environment," *Asian Journal of Control*, vol. 16, no. 6, pp. 1771–1778, 2014.

[17] M. Jafari, S. Sengupta, and H. M. La, "Adaptive flocking control of multiple unmanned ground vehicles by using a uav," in *International Symposium on Visual Computing*, pp. 628–637, Springer, 2015.

[18] S. Ghapani, J. Mei, W. Ren, and Y. Song, "Fully distributed flocking with a moving leader for lagrange networks with parametric uncertainties," *Automatica*, vol. 67, pp. 67–76, 2016.

[19] Q. Zhang, P. Li, Z. Yang, and Z. Chen, "Adaptive flocking of non-linear

multi-agents systems with uncertain parameters," *IET Control Theory & Applications*, vol. 9, no. 3, pp. 351–357, 2014.

[20] Y. Dong and J. Huang, "Flocking with connectivity preservation of multiple double integrator systems subject to external disturbances by a distributed control law," *Automatica*, vol. 55, pp. 197–203, 2015.

[21] Z. A. S. Dashti, M. Gholami, M. Jafari, M. A. Shoorehdeli, and M. Teshnehlab, "Speed control of a digital servo system using parallel distributed compensation controller and neural adaptive controller," in *Fuzzy Systems (IFSC), 2013 13th Iranian Conference on*, pp. 1–6, IEEE, 2013.

[22] Z. A. S. Dashti, M. Gholami, M. Jafari, and M. A. Shoorehdeli, "Neural—adaptive control based on backstepping and feedback linearization for electro hydraulic servo system," in *Intelligent Systems (ICIS), 2014 Iranian Conference on*, pp. 1–6, IEEE, 2014.

[23] M. Hajimani, Z. A. S. Dashti, M. Gholami, M. Jafari, and M. A. Shoorehdeli, "Neural adaptive controller for magnetic levitation system," in *Intelligent Systems (ICIS), 2014 Iranian Conference on*, pp. 1–6, IEEE, 2014.

[24] A. Hasanpour, P. Daemy, M. Aghazamani, K. Alipour, and M. Jafari, "Kinematic analysis of darwin's humanoid robot," in *Control, Instrumentation, and Automation (ICCIA), 2016 4th International Conference on*, pp. 356–361, IEEE, 2016.

[25] H. Samadi and M. Jafari, "Detecting assembled workpieces to sort according to color by mitsubishi rv-2aj robot and festo sorting station based on industrial networks," in *1st International Conference on New Research Achievements in Electrical and Computer Engineering (ICNRAECE)*, IEEE, 2016.

[26] A. Nourmohammadi, M. Jafari, and T. O. Zander, "A survey on unmanned aerial vehicle remote control using brain–computer interface," *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 4, pp. 337–348, 2018.

[27] Z. Li, N. Hovakimyan, and D. Stipanović, "Distributed multi-agent tracking and estimation with uncertain agent dynamics," in *American Control Conference (ACC), 2011*, pp. 2204–2209, IEEE, 2011.

[28] G. Hu, "Robust consensus tracking for an integrator-type multi-agent system with disturbances and unmodelled dynamics," *International journal of control*, vol. 84, no. 1, pp. 1–8, 2011.

[29] Z. Peng, D. Wang, W. Lan, G. Sun, and L. Yan, "Neural adaptive flocking control of networked underactuated autonomous surface vehicles in the presence of uncertain dynamics," in *Control Conference (CCC), 2012 31st Chinese*, pp. 2865–2870, IEEE, 2012.

[30] Z. Peng, D. Wang, H. H. Liu, and G. Sun, "Neural adaptive control for leader–follower flocking of networked nonholonomic agents with unknown nonlinear dynamics," *International Journal of Adaptive Control and Signal Processing*, vol. 28, no. 6, pp. 479–495, 2014.

[31] Q. Xiaowei and R. Guang, "Neural adaptive control for flocking of agents with unknown nonlinear dynamics," in *International Conference on Manufacturing Science and Engineering (ICMSE 2015)*, pp. 41–44, Atlantis Press, 2015.

[32] S.-M. Hung and S. N. Givigi, "A q-learning approach to flocking with uavs in a stochastic environment," *IEEE transactions on cybernetics*, vol. 47, no. 1, pp. 186–197, 2017.

[33] J. R. Lawton, R. W. Beard, and B. J. Young, "A decentralized approach to formation maneuvers," *IEEE Transactions on Robotics and Automation*, vol. 19, no. 6, pp. 933–941, 2003.

[34] B. Liu and H. Yu, "Flocking in multi-agent systems with a bounded control input," in *Chaos-Fractals Theories and Applications, 2009. IWCFTA'09. International Workshop on*, pp. 130–134, IEEE, 2009.

[35] A. Gasparri, G. Oriolo, A. Priolo, and G. Ulivi, "A swarm aggregation algorithm based on local interaction for multi-robot systems with actuator saturations," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 539–544, IEEE, 2012.

[36] A. Leccese, A. Gasparri, A. Priolo, G. Oriolo, and G. Ulivi, "A swarm aggregation algorithm based on local interaction with actuator saturations and integrated obstacle avoidance," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pp. 1865–1870, IEEE, 2013.

[37] P. Ke, S. Hou-Sheng, and Y. Yu-Pu, "Coordinated control of multi-agent systems with a varying-velocity leader and input saturation," *Communications in Theoretical Physics*, vol. 52, no. 3, pp. 449–456, 2009.

[38] D. V. Dimarogonas and K. H. Johansson, "Decentralized connectivity maintenance in mobile networks with bounded inputs," in *Robotics and Automa-*

tion, 2008. ICRA 2008. IEEE International Conference on, pp. 1507–1512, IEEE, 2008.

[39] Z. Chen, M.-C. Fan, and H.-T. Zhang, "How much control is enough for network connectivity preservation and collision avoidance?," *IEEE transactions on cybernetics*, vol. 45, no. 8, pp. 1647–1656, 2015.

[40] J. Moren and C. Balkenius, "A computational model of emotional learning in the amygdala," *From animals to animats*, vol. 6, pp. 115–124, 2000.

[41] C. Lucas, D. Shahmirzadi, and N. Sheikholeslami, "Introducing belbic: brain emotional learning based intelligent controller," *Intelligent Automation & Soft Computing*, vol. 10, no. 1, pp. 11–21, 2004.

[42] Z. Beheshti and S. Z. M. Hashim, "A review of emotional learning and it's utilization in control engineering," *Int. J. Adv. Soft Comput. Appl*, vol. 2, no. 2, pp. 191–208, 2010.

[43] M. Jafari, A. M. Shahri, and S. B. Shouraki, "Attitude control of a quadrotor using brain emotional learning based intelligent controller," in *Fuzzy Systems (IFSC), 2013 13th Iranian Conference on*, pp. 1–5, IEEE, 2013.

[44] E. Lotfi and M.-R. Akbarzadeh-T, "A winner-take-all approach to emotional neural networks with universal approximation property," *Information Sciences*, vol. 346, pp. 369–388, 2016.

[45] M. Jafari, "Controlling a quadrotor using brain emotional learning based intelligent controller," Master's thesis, Qazvin Islamic Azad University, 2012.

[46] M. Jafari, S. B. Shuraki, *et al.*, "Speed control of a digital servo system using brain emotional learning based intelligent controller," in *Power Electronics, Drive Systems and Technologies Conference (PEDSTC), 2013 4th*, pp. 311–314, IEEE, 2013.

[47] M. Jafari, A. M. Shahri, and S. H. Elyas, "Optimal tuning of brain emotional learning based intelligent controller using clonal selection algorithm," in *Computer and Knowledge Engineering (ICCKE), 2013 3th International eConference on*, pp. 30–34, IEEE, 2013.

[48] M. Jafari, R. Fehr, L. R. G. Carrillo, and H. Xu, "Brain emotional learning-based intelligent tracking control for unmanned aircraft systems with un-

certain system dynamics and disturbance," in *Unmanned Aircraft Systems (ICUAS), 2017 International Conference on*, pp. 1470–1475, IEEE, 2017.

[49] M. Jafari, H. Xu, and L. R. Garcia Carrillo, "A neurobiologically-inspired intelligent trajectory tracking control for unmanned aircraft systems with uncertain system dynamics and disturbance," *Transactions of the Institute of Measurement and Control*, 2018.

[50] M. Jafari and H. Xu, "Adaptive neural network based intelligent control for unmanned aerial systems with system uncertainties and disturbances," in *Unmanned Aircraft Systems (ICUAS), 2018 International Conference on*, IEEE, 2018.

[51] M. Jafari, H. Xu, and L. R. G. Carrillo, "Brain emotional learning-based intelligent controller for flocking of multi-agent systems," in *American Control Conference (ACC), 2017*, pp. 1996–2001, IEEE, 2017.

[52] R. Fehr, K. Boles, M. Jafari, H. Xu, and L. R. G. Carrillo, "A low-computation distributed connectivity control for coordinated multi-uas," in *Unmanned Aircraft Systems (ICUAS), 2017 International Conference on*, pp. 1181–1188, IEEE, 2017.

[53] M. Jafari, R. Fehr, L. R. G. Carrillo, E. S. E. Quesada, and H. Xu, "Implementation of brain emotional learning-based intelligent controller for flocking of multi-agent systems," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 6934–6939, 2017.

[54] M. Jafari and H. Xu, "A biologically-inspired intelligent controller for distributed velocity control of multiple electro-hydraulic servo-systems," in *Computational Intelligence (SSCI), 2017 IEEE Symposium Series on*, pp. 1–7, IEEE, 2017.

[55] M. Jafari, H. Xu, and L. R. Garcia Carrillo, "A biologically–inspired reinforcement learning based intelligent distributed flocking control for multi–agent systems in presence of uncertain system and dynamic environment," *IFAC Journal of Systems and Control*.

[56] M. Jafari and H. Xu, "A biologically-inspired distributed resilient flocking control for multi-agent system with uncertain dynamics and unknown disturbances," in *Resilience Week (RWS), 2017*, pp. 71–76, IEEE, 2017.

[57] M. Jafari and H. Xu, "A biologically–inspired distributed fault tolerant flock-

ing control for multi–agent system in presence of uncertain dynamics and unknown disturbance," *Engineering applications of artificial intelligence*.

[58] M. Jafari and H. Xu, "A biologically-inspired distributed intelligent flocking control for networked multi-uas with uncertain network imperfections," in *Unmanned Aircraft Systems (ICUAS), 2018 International Conference on*, IEEE, 2018.

[59] M. Jafari, H. Xu, and L. R. Garcia Carrillo, "Brain emotional learning-based path planning and intelligent control co-design for unmanned aerial vehicle in presence of system uncertainties and dynamic environment," in *Computational Intelligence (SSCI), 2018 IEEE Symposium Series on*, IEEE, 2018.

[60] M. Jafari, V. Sarfi, A. Ghasemkhani, H. Livani, L. Yang, and H. Xu, "Biologically–inspired adaptive intelligent secondary control for microgrids under cyber imperfections," *IET Cyber-Physical Systems: Theory & Applications*.

[61] M. Jafari and H. Xu, "A game theoretic based biologically-inspired distributed intelligent flocking control for multi-uav systems with network imperfections," in *Computational Intelligence (SSCI), 2018 IEEE Symposium Series on*, IEEE, 2018.

[62] L. R. G. Carrillo, A. E. D. López, R. Lozano, and C. Pégard, *Quad rotorcraft control: vision-based hovering and navigation*. Springer Science & Business Media, 2012.

[63] F. Kendoul, "Survey of advances in guidance, navigation, and control of unmanned rotorcraft systems," *Journal of Field Robotics*, vol. 29, no. 2, pp. 315–378, 2012.

[64] H. J. Asl and H. Bolandi, "Robust vision-based control of an underactuated flying robot tracking a moving target," *Transactions of the Institute of Measurement and Control*, vol. 36, no. 3, pp. 411–424, 2014.

[65] H. Du, W. Zhu, G. Wen, Z. Duan, and J. Lü, "Distributed formation control of multiple quadrotor aircraft based on nonsmooth consensus algorithms," *IEEE Transactions on Cybernetics*, 2017.

[66] C. Coza, C. Nicol, C. Macnab, and A. Ramirez-Serrano, "Adaptive fuzzy control for a quadrotor helicopter robust to wind buffeting," *Journal of Intelligent*

& Fuzzy Systems: Applications in Engineering and Technology, vol. 22, no. 5, 6, pp. 267–283, 2011.

[67] Y. Zhang, B. Xu, and H. Li, "Adaptive neural control of a quadrotor helicopter with extreme learning machine," in *Proceedings of ELM-2014 Volume 2*, pp. 125–134, Springer, 2015.

[68] S. Islam, M. Faraz, R. Ashour, G. Cai, J. Dias, and L. Seneviratne, "Adaptive sliding mode control design for quadrotor unmanned aerial vehicle," in *Unmanned Aircraft Systems (ICUAS), 2015 International Conference on*, pp. 34–39, IEEE, 2015.

[69] H. Bou-Ammar, H. Voos, and W. Ertel, "Controller design for quadrotor uavs using reinforcement learning," in *Control Applications (CCA), 2010 IEEE International Conference on*, pp. 2130–2135, IEEE, 2010.

[70] B. Xian, C. Diao, B. Zhao, and Y. Zhang, "Nonlinear robust output feedback tracking control of a quadrotor uav using quaternion representation," *Nonlinear Dynamics*, vol. 79, no. 4, pp. 2735–2752, 2015.

[71] H. Liu, D. Li, J. Xi, and Y. Zhong, "Robust attitude controller design for miniature quadrotors," *International Journal of Robust and Nonlinear Control*, 2015.

[72] H. Wang and M. Chen, "Trajectory tracking control for an indoor quadrotor uav based on the disturbance observer," *Transactions of the Institute of Measurement and Control*, vol. 38, no. 6, pp. 675–692, 2016.

[73] M. Li, Z. Zuo, H. Liu, C. Liu, and B. Zhu, "Adaptive fault tolerant control for trajectory tracking of a quadrotor helicopter," *Transactions of the Institute of Measurement and Control*, p. 0142331217728568, 2017.

[74] A. Vargas-Clara and S. Redkar, "Unmanned ground vehicle navigation using brain emotional learning based intelligent controller (belbic)," *Smart Science*, vol. 3, no. 1, pp. 10–15, 2015.

[75] M.-H. Khooban and R. Javidan, "A novel control strategy for dvr: Optimal bi-objective structure emotional learning," *International Journal of Electrical Power & Energy Systems*, vol. 83, pp. 259–269, 2016.

[76] D. Shahmirzadi, *Computational modeling of the brain limbic system and its application in control engineering*. PhD thesis, Texas A&M University, 2005.

[77] T. Bresciani, *Modelling, identification and control of a quadrotor helicopter*. Department of Automatic Control, Lund University, 2008.

[78] C. Balkenius and J. Morén, "Emotional learning: A computational model of the amygdala," *Cybernetics & Systems*, vol. 32, no. 6, pp. 611–636, 2001.

[79] H. T. Dorrah, A. M. El-Garhy, and M. E. El-Shimy, "Pso-belbic scheme for two-coupled distillation column process," *Journal of advanced Research*, vol. 2, no. 1, pp. 73–83, 2011.

[80] S. Jafarzadeh, M. J. Motlagh, M. Barkhordari, and R. Mirheidari, "A new lyapunov based algorithm for tuning belbic controllers for a group of linear systems," in *Control and Automation, 2008 16th Mediterranean Conference on*, pp. 593–595, IEEE, 2008.

[81] N. Garmsiri and F. Najafi, "Fuzzy tuning of brain emotional learning based intelligent controllers," in *Intelligent Control and Automation (WCICA), 2010 8th World Congress on*, pp. 5296–5301, July 2010.

[82] H. Zaki, M. Unel, and Y. Yildiz, "Trajectory control of a quadrotor using a control allocation approach," in *Unmanned Aircraft Systems (ICUAS), 2017 International Conference on*, pp. 533–539, IEEE, 2017.

[83] F. M. Palacios, E. S. E. Quesada, G. Sanahuja, S. Salazar, O. G. Salazar, and L. R. G. Carrillo, "Test bed for applications of heterogeneous unmanned vehicles," *International Journal of Advanced Robotic Systems*, vol. 14, no. 1, 2017.

[84] H. Rouhani, M. Jalili, B. N. Araabi, W. Eppler, and C. Lucas, "Brain emotional learning based intelligent controller applied to neurofuzzy model of micro-heat exchanger," *Expert Systems with Applications*, vol. 32, no. 3, pp. 911–918, 2007.

[85] M. A. Sharbafi, C. Lucas, and R. Daneshvar, "Motion control of omni-directional three-wheel robots by brain-emotional-learning-based intelligent controller," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 6, pp. 630–638, 2010.

[86] M. A. Rahman, R. M. Milasi, C. Lucas, B. N. Araabi, and T. S. Radwan, "Implementation of emotional controller for interior permanent-magnet synchronous motor drive," *IEEE Transactions on Industry Applications*, vol. 44, no. 5, pp. 1466–1476, 2008.

[87] M. R. Jamali, M. Dehyadegari, A. Arami, C. Lucas, and Z. Navabi, "Real-time embedded emotional controller," *Neural Computing and Applications*, vol. 19, no. 1, pp. 13–19, 2010.

[88] Y. Cao, W. Yu, W. Ren, and G. Chen, "An overview of recent progress in the study of distributed multi-agent coordination," *IEEE Transactions on Industrial informatics*, vol. 9, no. 1, pp. 427–438, 2013.

[89] G. D. L. Torre and T. Yucelen, "Adaptive architectures for resilient control of networked multiagent systems in the presence of misbehaving agents," *International Journal of Control*, pp. 1–13, 2017.

[90] K. Saulnier, D. Saldana, A. Prorok, G. J. Pappas, and V. Kumar, "Resilient flocking for mobile robot teams," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 1039–1046, 2017.

[91] M. Jafari, I. Dehghan Ebrahimi, A. Torabi, and M. Karimi, "Using hopfield neural network for password authentication," in *Regional Conference on Computer Engineering, Knowledge & Information Technology*, 2012.

[92] H. Karimi, B. Moshiri, B. Lohmann, and P. J. Maralani, "Haar wavelet-based approach for optimal control of second-order linear systems in time domain," *Journal of Dynamical and Control Systems*, vol. 11, no. 2, pp. 237–252, 2005.

[93] H. R. Karimi and H. Gao, "Mixed H_2/H_∞ output-feedback control of second-order neutral systems with time-varying state and input delays," *ISA transactions*, vol. 47, no. 3, pp. 311–324, 2008.

[94] M. Jafari, V. Sarfi, A. Ghasemkhani, H. Livani, L. Yang, H. Xu, and R. Koosha, "Adaptive neural network based intelligent secondary control for microgrids," in *Texas Power and Energy Conference (TPEC), 2018 IEEE*, pp. 1–6, IEEE, 2018.

[95] A. Ghasemkhani and L. Yang, "Reinforcement learning based pricing for demand response," in *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, IEEE, 2018.

[96] E. Daryabeigi, N. R. Abjadi, and G. Arab Markadeh, "Automatic speed control of an asymmetrical six-phase induction motor using emotional controller (belbic)," *Journal of Intelligent & Fuzzy Systems*, vol. 26, no. 4, pp. 1879–1892, 2014.

[97] R. Farhangi, M. Boroushaki, and S. H. Hosseini, "Load–frequency control of interconnected power system using emotional learning-based intelligent controller," *International Journal of Electrical Power & Energy Systems*, vol. 36, no. 1, pp. 76–83, 2012.

[98] J. Moren, *Emotion and learning-A computational model of the amygdala*, vol. 93. Lund University Cognitive Science, 2002.

[99] H. R. Karimi, "Robust synchronization and fault detection of uncertain master-slave systems with mixed time-varying delays and nonlinear perturbations," *International Journal of Control, Automation and Systems*, vol. 9, no. 4, p. 671, 2011.

[100] R. Sakthivel, P. Selvaraj, Y. Lim, and H. Karimi, "Adaptive reliable output tracking of networked control systems against actuator faults," *Journal of the Franklin Institute*, vol. 354, no. 9, pp. 3813–3837, 2017.

[101] Y. Tang, X. Xing, H. R. Karimi, L. Kocarev, and J. Kurths, "Tracking control of networked multi-agent systems under new characterizations of impulses and its applications in robotic systems," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 2, pp. 1299–1307, 2016.

[102] H. Liu, H. R. Karimi, S. Du, W. Xia, and C. Zhong, "Leader-following consensus of discrete-time multiagent systems with time-varying delay based on large delay theory," *Information Sciences*, vol. 417, pp. 236–246, 2017.

[103] D. Zhang, Z. Xu, H. R. Karimi, Q.-G. Wang, and L. Yu, "Distributed H_∞ output-feedback control for consensus of heterogeneous linear multi-agent systems with aperiodic sampled-data communications," *IEEE Transactions on Industrial Electronics*, 2017.

[104] W. Xing, Y. Zhao, and H. R. Karimi, "Convergence analysis on multi-auv systems with leader-follower architecture," *IEEE Access*, vol. 5, pp. 853–868, 2017.

[105] S. Arora and B. Barak, *Computational complexity: a modern approach*. Cambridge University Press, 2009.

[106] E. Lotfi and M.-R. Akbarzadeh-T, "Adaptive brain emotional decayed learning for online prediction of geomagnetic activity indices," *Neurocomputing*, vol. 126, pp. 188–196, 2014.

[107] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on automatic control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[108] N. Chopra and M. W. Spong, "Output synchronization of nonlinear systems with time delay in communication," in *Decision and Control, 2006 45th IEEE Conference on*, pp. 4986–4992, IEEE, 2006.

[109] S. Kar and J. M. Moura, "Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 355–369, 2009.

[110] Y. Liu, D. W. Ho, and Z. Wang, "A new framework for consensus for discrete-time directed networks of multi-agents with distributed delays," *International Journal of Control*, vol. 85, no. 11, pp. 1755–1765, 2012.

[111] Z. Wan, "Flocking for multi-agent dynamical systems," Master's thesis, University of Waterloo, 2012.

[112] Z. Yang, Q. Zhang, Z. Jiang, and Z. Chen, "Flocking of multi-agents with time delay," *International journal of systems science*, vol. 43, no. 11, pp. 2125–2134, 2012.

[113] Q. Zhang, P. Li, Z. Yang, and Z. Chen, "Distance constrained based adaptive flocking control for multiagent networks with time delay," *Mathematical Problems in Engineering*, vol. 2015, 2015.

[114] A. R. Mehrabian and K. Khorasani, "Distributed formation recovery control of heterogeneous multiagent euler–lagrange systems subject to network switching and diagnostic imperfections," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 6, pp. 2158–2166, 2016.

[115] Y. Cao and T. Oguchi, "Coordinated control of mobile robots with delay compensation based on synchronization," in *Sensing and Control for Autonomous Vehicles*, pp. 495–514, Springer, 2017.

[116] I. Niazazari and H. Livani, "A pmu-data-driven disruptive event classification in distribution systems," *Electric Power Systems Research*, vol. 157, pp. 251–260, 2018.

[117] J.-W. Kim, C.-Y. Oh, J.-W. Chung, and K.-H. Kim, "Brain emotional limbic-

based intelligent controller design for control of a haptic device," *International Journal of Automation and Control*, vol. 11, no. 4, pp. 358–371, 2017.

[118] C. Rizzi, C. G. Johnson, F. Fabris, and P. A. Vargas, "A situation-aware fear learning (safel) model for robots," *Neurocomputing*, vol. 221, pp. 32–47, 2017.

[119] J. R. Marden and J. S. Shamma, "Game theory and control," *Annual Review of Control, Robotics, and Autonomous Systems*, no. 0, 2017.

[120] W. Lin, Z. Qu, and M. A. Simaan, "Distributed game strategy design with application to multi-agent formation control," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pp. 433–438, IEEE, 2014.

[121] Y. Ho, A. Bryson, and S. Baron, "Differential games and optimal pursuit-evasion strategies," *IEEE Transactions on Automatic Control*, vol. 10, no. 4, pp. 385–389, 1965.

[122] S. M. LaValle, *Planning algorithms*. Cambridge university press, 2006.

[123] C. Goerzen, Z. Kong, and B. Mettler, "A survey of motion planning algorithms from the perspective of autonomous uav guidance," *Journal of Intelligent and Robotic Systems*, vol. 57, no. 1, pp. 65–100, 2010.

[124] C. Qixin, H. Yanwen, and Z. Jingliang, "An evolutionary artificial potential field algorithm for dynamic path planning of mobile robot," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pp. 3331–3336, IEEE, 2006.

[125] W. Afzal and A. A. Masoud, "Harmonic potential-based communication-aware navigation of mobile agents in cluttered spaces," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pp. 5146–5151, IEEE, 2016.

[126] Y. Shoukry, P. Nuzzo, I. Saha, A. L. Sangiovanni-Vincentelli, S. A. Seshia, G. J. Pappas, and P. Tabuada, "Scalable lazy smt-based motion planning," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pp. 6683–6688, IEEE, 2016.

[127] W. Yao, N. Wan, and N. Qi, "Hierarchical path generation for distributed mission planning of uavs," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pp. 1681–1686, IEEE, 2016.

[128] J. Moreau, P. Melchior, S. Victor, F. Aioun, and F. Guillemard, "Path planning with fractional potential fields for autonomous vehicles," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 14533–14538, 2017.

[129] M. Rafieisakhaei, S. Chakravorty, and P. Kumar, "A near-optimal decoupling principle for nonlinear stochastic systems arising in robotic path planning and control," in *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on*, pp. 1–6, IEEE, 2017.

[130] Y. Mei, G. Tan, and Z. Liu, "An improved brain-inspired emotional learning algorithm for fast classification," *Algorithms*, vol. 10, no. 2, p. 70, 2017.

[131] M. A. A.-A. El-Garhy, R. Mubarak, and M. El-Bably, "Improving maximum power point tracking of partially shaded photovoltaic system by using ipso-belbic," *Journal of Instrumentation*, vol. 12, no. 08, p. P08012, 2017.

[132] C.-M. Lin and C.-C. Chung, "Fuzzy brain emotional learning control system design for nonlinear systems," *International Journal of Fuzzy Systems*, vol. 17, no. 2, pp. 117–128, 2015.