CENTRO DE INVESTIGACIÓN Y ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD ZACATENCO
DEPARTAMENTO DE CONTROL AUTOMÁTICO

Realización Robusta de un Controlador Óptimo Basado en Programación Dinámica Neuronal Aproximada

Tesis que Presenta

Mariana Felisa Ballesteros Escamilla

Para obtener el grado de
Doctor en Ciencias

En la especialidad de
Control Automático

Directores de tesis:

Dr. Alexander Poznyak
Dr. Jorge Isaac Chairez Oria

Ciudad de México

12 de Febrero del 2020

CENTER FOR RESEARCH AND ADVANCED STUDIES OF THE NATIONAL POLYTECHNIC INSTITUTE

ZACATENCO CAMPUS

AUTOMATIC CONTROL DEPARTMENT

Robust Realization of Optimal Control Based on Approximate Neural Dynamic Programming

Thesis presented by

Mariana Felisa Ballesteros Escamilla

To obtain the Degree of

Doctor of Philosophy

in

Automatic Control Science

Thesis supervised by:

Dr. Alexander Poznyak
Dr. Jorge Isaac Chairez Oria

Mexico City

12th February 2020

“Deja que todo te suceda: la belleza y el terror.

Sólo sigue adelante.

Ningún sentimiento es definitivo”

Rilke, R. M. El libro de horas. [1]

Para mi mamá, quien me ha enseñado que siempre debo de continuar.

Jurado Asignado

Investigadores CINVESTAV

- Dr. Moisés Bonilla Estrada
- Dr. Fernando Castaños Luna
- Dr. Wen Yu

Investigadores Externos

- Dr. Iván de Jesús Salgado Ramos

Director interno

- Dr. Alexander Poznyak

Director externo

- Dr. Jorge Isaac Chairez Oria

Agradecimientos

“Agradezcamos a las personas que nos dan felicidad,
son los adorables jardineros que hacen florecer nuestras almas”.
Marcel Proust. Los placeres y los días. [2]

Principalmente a ti, todo te lo debo a ti. Gracias por tu apoyo, por todos los consejos, por confiar en mí y dejarme soñar, por todo lo que me has enseñado y sobre todo por tu amor incondicional.

Te amo, mamá.

Haces que mi vida sea más bonita, gracias por tu amor y por compartir conmigo tantos momentos y experiencias. Gracias también por toda la ayuda, por los ánimos que siempre me das y por estar conmigo a pesar de los tiempos difíciles.

David Cruz.

Gracias a toda mi familia, la familia Ballesteros. Los quiero mucho y sé que cuento con ustedes, así como ustedes cuentan conmigo.

Es verdad que la familia también se elige. A todos mis amigos, los quiero muchísimo, gracias por el apoyo, por la alegría que traen a mi vida, por todas las vivencias, las risas, las pláticas y hasta los momentos difíciles que hemos compartido.

Gracias a todas las buenas personas que de alguna forma me apoyaron para que pudiera continuar estudiando, no hubiera podido llegar al doctorado si en su momento no hubieran estado allí para darme una mano. Siempre las tengo presente y las recuerdo con mucho cariño.

A mi director, Dr. Isaac, por todas sus enseñanzas y el tiempo compartido. Gracias por ser un excelente director y profesor, por la confianza que me ha tenido para permitirme ser su alumna desde la licenciatura hasta el doctorado. Es para mí una persona admirable que siempre muestra esfuerzo y dedicación en todos los proyectos, realmente es un ejemplo. Gracias también por sus

mejores enseñanzas que son la humildad y la amabilidad, porque siempre mantiene la puerta abierta a todos nosotros, sus estudiantes. Usted me motiva a siempre seguir aprendiendo.

Al Dr. Alexander Poznyak, muchas gracias por la sabiduria compartida y el tiempo dedicado. Gracias por aceptar ser mi director de tesis, por el apoyo, por todos los consejos y las enseñanzas. Es usted una persona admirable.

A los doctores Andrey Polyakov y Denis Effimov, por aceptarme en la estancia de investigación en INRIA. Gracias por compartir sus conocimientos, por el apoyo brindado y por todos los comentarios y consejos para mejorar el trabajo.

Al Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, por todas las facilidades, los apoyos y la educación durante mi doctorado.

Al CONACyT, por la beca otorgada (CVU 550803) para poder llevar a cabo mis estudios de posgrado y por el apoyo para la realización de la estancia de investigación.

Abstract

The aim of this thesis was to design an artificial neural network (ANN) based sub-optimal controller for a finite horizon optimization problem. The class of systems considered in this thesis were continuous systems with parametric uncertainties and bounded external perturbations. The control design approach considered the uncertain part of the system by proposing a min-max problem. The min-max problem was tackled using a neural dynamic programming solution, this approach was selected due to the universal approximation capabilities of ANNs.

An ANN structure approximates the Value function (VF) used in the solution of the Hamilton-Jacobi-Bellman (HJB) equation. The network topology was proposed to satisfy the characteristics of the VF including positiveness and continuity. The explicit adaptive laws, that adjusted the weights in the ANN, were obtained directly from the HJB approximated solution. The stability analysis based on the Lyapunov theory yields to confirm that the approximated VF can be used as a Lyapunov function candidate to confirm the practical stability of the origin in the stabilization problem.

Two cases of the studied class of systems were considered. First, a system with additive uncertain term, then, in the second case, an uncertain multiplicative element with the input was added.

The approximation capabilities of ANNs were also used to design an ANN identifier. The identifier consisted on a differential neural network (DNN) structure and the proposed learning laws ensures the ultimate boundedness and the input to state stability of the weights and the identification error. The study on the identifier extends the approximation capabilities of ANNs for a class of homogeneous nonlinear systems.

Homogeneous systems has a symmetry-like property under which an object remains consistent with respect to a certain scaling or dilation. Homogeneity allows local properties of vector fields to be extended globally.

The DNN-identifier obtains an approximated model for the class of uncertain nonlinear homogeneous systems, this solution can be used for the design of controllers such as the optimal control problem presented in this thesis. In addition, the result of the extension of the ANN approximation capabilities could be helpful to obtain approximated VF structures with the homogeneous properties.

Resumen

El objetivo de esta tesis fue diseñar un controlador subóptimo utilizando redes neuronales artificiales para obtener la solución de un problema de optimización de horizonte finito. La clase de sistemas considerados en el problema son sistemas continuos con incertidumbres paramétricas y perturbaciones externas acotadas.

El enfoque de diseño de control considera la parte incierta del sistema al proponer un problema min-máx. Para la solución del problema min-máx se utilizó la técnica de programación dinámica neuronal, se seleccionó esta técnica por las propiedades de aproximación universal que poseen las redes neuronales artificiales.

Se propuso el uso de una estructura especial de red neuronal artificial para aproximar la solución o la función Valor para la ecuación de Hamilton-Jacobi-Bellman. La topología de red se propuso para satisfacer las características de la función Valor como positividad y continuidad. La ley adaptable para ajustar los pesos de la red neuronal artificial se obtuvo directamente de la solución aproximada de la ecuación de Hamilton-Jacobi-Bellman. Se realizó el análisis de estabilidad con base en la teoría de Lyapunov, esto permitió confirmar que la función Valor aproximada se puede utilizar como una función candidata de Lyapunov para confirmar la estabilidad práctica del origen en el problema de estabilización.

Se consideraron dos casos en la clase de sistemas estudiados. Primero, un sistema con un término desconocido de forma aditiva, este término representa la incertidumbre o perturbación acotada, después, en el segundo caso, se añadió un elemento desconocido asociado a la entrada en forma multiplicativa.

Las propiedades de aproximación de las redes neuronales también fueron utilizadas para el diseño de un identificador basado en redes neuronales. El identificador consiste en una estructura de red neuronal diferencial cuyas leyes de aprendizaje se proponen de tal forma que se asegure que el error de identificación y los pesos de ajuste están últimamente acotados, probando también estabilidad de entrada-estado. El estudio del identificador extiende las propiedades de aproximación de las redes neuronales artificiales para una clase de sistemas no lineales homogéneos.

Los sistemas homogéneos poseen un tipo de simetría, esta propiedad permite que un objeto siga siendo consistente respecto a cierta escala o dilación. La homogeneidad en los campos vectoriales permite que propiedades que son probadas localmente puedan ser extendidas de forma global.

El identificador basado en redes neuronales diferenciales obtiene una aproximación del modelo para la clase de sistemas no lineales homogéneos, esta solución puede ser utilizada para el diseño de controladores, por ejemplo, para el problema de optimización presentado en esta tesis. Además, el resultado de la extensión de las propiedades de aproximación de las redes neuronales artificiales puede ser útil para obtener soluciones aproximadas de la función Valor contando con las propiedades homogéneas.

Contents

List of Figures

List of Tables

Acronyms

- **A**

  ANN Artificial Neural Network
- **D**

  DNN Differential Neural Networks

  DP Dynamic Programming
- **H**

  HJB Hamilton-Jacobi-Bellman
- **I**

  ISS Input to the state stability
- **L**

  LQR Linear Quadratic Regulator
- **M**

  MP Maximum Principle
- **N**

  NDP Neuro Dynamic Programming
- **O**

  OC Optimal Control

  OCP Optimal Control Problem

  ODE Ordinary Differential Equation

- **P**

  PE Persistent Excitation

  PDE Partial Differential Equation

  PO Principle of Optimality

- **R**

  RL Reinforcement Learning

- **V**

  VF Value Function

Notation

The following notation is used through this thesis:

- I_n is the identity matrix of $n \times n$ dimension.
- $\mathbb{R}$ is the set of real numbers. $\mathbb{R}^n \setminus \{0\}$ denotes the difference of the set $\mathbb{R}$ and $\{0\}$, i.e., all real numbers except zero. $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$ is the set of the positive real numbers, including zero. $\mathbb{R}^n$ is the set of column vectors of dimension $n \times 1$.
- $A^\top$ denotes the transpose operator over the matrix $A \in \mathbb{R}^{n \times m}$ (or vector), i.e., the i-th row, j-th column element of $A^\top$ is the j-th row, i-th column element of A.
- $\|\cdot\|$ represents the Euclidean norm. $\|z\|_H = \sqrt{z^\top H z}$ is the weighted norm for $z \in \mathbb{R}^n$ and $H \in \mathbb{R}^{n \times n}$, $H = H^\top$, $H > 0$. $\|M\|_F = \sqrt{\mathrm{tr}\left\{M^\top M\right\}}$ is the Frobenius norm.
- $\mathrm{tr}\{H\} = \sum_{i=1}^{n} H_{i,i}$ is the trace operator and H_{ii} is the element of matrix $H \in \mathbb{R}^{n \times n}$ in row i and column i.
- $\nabla F(x)$ is the operator representing the gradient or the partial derivative respect $x \in \mathbb{R}^n$, $\frac{\partial}{\partial x} F(x)$ of the function $F : \mathbb{R}^n \to \mathbb{R}$.
- ν^* represents that such value of ν is an extremal or optimal (minimum or maximum).
- $u_{[a,b)}$ indicates that the signal u is restricted to the interval $[a, b)$.
- $V \in C^1$ the V function is continuously differentiable.
- $o(\alpha)$ represents the o-small ("high-order terms"), which go faster to zero than α goes, i.e., $\lim_{\alpha \to 0} \frac{o(\alpha)}{\alpha} = 0$.
- $\langle x, y \rangle$ represents the inner product, $x \in \mathbb{R}^n$, $z \in \mathbb{R}^n$, $\langle x, y \rangle = x^\top y$.

- The Kronecker product is denoted by $\otimes$.
- For any $\vartheta \in \mathbb{R}_+$ and $\forall x \in \mathbb{R}$ we set $\lceil x \rfloor^{\vartheta} = \mathrm{sign}(x)|x|^{\vartheta}$.
- For $M \in \mathbb{R}^{m \times n}$, $M = [m_1, m_2, \ldots, m_n]$, we set $\mathrm{vec}(M) = [m_1^\top, m_2^\top, \ldots, m_n^\top]^\top$.
- A class-$\mathcal{K}$ function is a continuous function $\phi : [0, a) \to [0, \infty)$, $a \in \mathbb{R}_+$, if it is strictly increasing and $\phi(0) = 0$.

Chapter 1

Introduction

This chapter presents the general introduction, the motivation of this research, the main objectives and the formal definitions of the three main concepts on which this work is based, that is, the principle of optimality, dynamic programming and the approximation capabilities of artificial neural networks. In addition, the state of art and the description of the thesis structure are presented at the end of this chapter.

Numerous dynamical systems are represented with a mathematical model, for example, mechanical, biological, financial, electronic, among other class of plants are described using ordinary differential equations (ODEs) [3,4]. The mathematical representation of a dynamical system, among other applications, serves to design the input signals. It is possible to drive the system from one state to another using a determinate input signal or signals[1], moreover, there is not only one possible control input to regulate the system [6].

One of the main objectives of automatic control theory is the regulation of different systems by designing appropriate input signals. For some applications the interest in the design of the control is to find the one which represents the "*best*" way to do a task according to a given criterion, for example, the time to go from one state to another, the energy of the input signal

[1] Considering the controllability property of the system [3, 5].

or the convergence of the state to an equilibrium. In real applications, this could be traduced as the less waste of fuel in a vehicle, the minimum error in the trajectory tracking problem for a robot, the maximum thrust developed by a missile, etc. This class of control requires the exact mathematical model and the input to achieve the minimization or maximization is known as *optimal control* (OC), while the criterion is called *cost functional* [7].

The base of the OC theory started with the study of the calculus of variations [8] three centuries ago and in the decade of 1960 the two principal approaches in modern control theory to solve an optimal control problem (OCP) arised, the Maximum Principle (MP) of Pontryagin [9] and the Bellman's Principle of Optimality (PO) [10].

The theoretical framework for linear systems is well established and both theories (the MP and the PO) give the necessary and sufficient conditions, however, in the case of nonlinear systems, finding the exact analytic solution is a difficult task. The design of the OC needs the exact model and parameters of the system, however, for real applications, to get the exact mathematical model may seem to be unfeasible. In addition, it is important to consider the model uncertainties and the presence of perturbations. To get an approximate linear model is a possible solution to obtain an OC, however, this is a local solution, which is valid in a small region around the origin. Moreover, the OC solution for the linearized system does not consider nor eliminate the effects of the uncertainties and perturbations.

For the case of nonlinear plants, the MP approach states the necessary conditions for optimality and leads to a nonlinear two-point boundary value problem [11], meanwhile, the Dynamic Programming (DP) approach, which is derived from the PO, states sufficient and necessary conditions[2]. The DP approach uses the definition of Value function (VF) and leads to finding the solution of a partial differential equation (PDE) [13].

The MP approach gives open-loop solutions. On the other hand, with the DP approach, state feedback control laws are obtained. This is an important feature from the practical implementation point of view, nonetheless, the feedback solution cannot be implemented unless the PDE can be solved [12]. The PDE derived from the DP approach is known as the Hamilton-Jacobi-Bellman (HJB) equation. In general it is a nonlinear first order with respect to time and the state as well

[2]Under smoothness assumptions on the VF, otherwise, this approach can only state sufficient conditions [12].

PDE and finding the solution is a challenging task.

The interest of this thesis is the design of a feedback OC for a class of continuous dynamical affected by modeling uncertainties or external perturbations. This approach uses the PO to obtain a closed loop structure. Moreover, to find the solution of the PDE (HJB equation) is proposed an approximation based on artificial neural networks (ANNs).

1.1 Motivation

The PO drives to the DP approach to solve OCPs with closed loop solutions and this approach is used for finite horizon optimization problems. Some applications of OC in real systems require closed loop solutions in finite time [7]. Nonetheless the general theory of DP is well studied, solving the finite horizon OCP for nonlinear systems with a DP approach is a difficult task due to the complexity of finding the solution for the HJB equation. In addition, the necessity of having the exact mathematical model of the system can not be neglected.

According with the DP literature review, there exist numerous works related with finding the solution for the HJB equation, for example viscosity solutions [14], discrete approximations [15], model predictive control [16], reinforcement learning (RL) [17] and ANNs [18]. On one hand, this last method is useful because of the universal approximation properties of ANNs [19]. On the other hand, the problem of estimating the exact mathematical representation of the plant or system to be optimally controlled is another motivation. This problem has been solved by using mixed controllers with classical robust techniques or by modifying the cost functionals by adding the uncertain elements [20, 21].

Two main ideas are proposed in this work: the first one consists on including the effect of the unknown bounded elements in the Hamiltonian and to set a min-max problem to obtain a robust-like solution. This approach has been used to solve OCPs for multimodel systems[3]. The second proposition is the approximated solution for the max-min HJB equation based on ANN structures. ANNs are used to approximate solutions for PDEs. The use of ANNs in DP (to ap-

[3] In [22], the OCP for the finite case (miltimodel systems) was set with the min-max approach using Lagrange multipliers and solved using the MP.

proximate the HJB equation) is known as Neuro-Dynamic Programming (NDP) [23], this method has demonstrated sub-optimal results for different class of systems, this is a big motivation of the thesis.

1.2 Objectives

1.2.1 General objective

The main objective of the thesis is to present an ANN feedback control solution for the finite horizon OCP for a class of continuous dynamic system affected by bounded perturbations or model uncertainties.

1.2.2 Particular objectives

Within the scope of this investigation work, the following particular objectives are stated:

- To set a min-max problem for continuous systems with bounded perturbations or model uncertainties in order to solve the finite horizon OCP with the DP approach.
- To derive the max-min HJB equation for the class of systems using the theory of the PO, setting the sufficient and necessary conditions for the optimality with the min-max proposition[4].
- To find the optimal arguments for the Hamiltonian, i.e., to minimize the Hamiltonian with respect to the unknown elements and to maximize the Hamiltonian to obtain the OC expression.
- To approximate the solution for the max-min HJB equation using ANN structures, which consider the characteristics of such solution[5]. In addition, to study the impact of the approximation quality using the ANN structure in the optimality.

[4] Under smoothness assumption on the max-min HJB equation solution.

[5] Positiveness, continuity and differentiability, this properties are also used for the stability and optimality analysis.

- To obtain the learning laws using the max-min HJB equation and to propose a numerical solution to adjust the free parameters or initial conditions of the parameters for the ANN structure.
- To use the approximated solution to compute the OC and to study the closed loop system stability with the Lyapunov theory. Finally, to evaluate the designed control with different numerical simulations.

1.3 Theoretical Framework

The objective of OC theory is to determine the control signals that will cause a process to satisfy the physical constraints and at the same time optimize (minimize or maximize) some performance criterion [9]. The formulation of an OCP needs three main elements:

(i) The mathematical description (or model) of the process to be controlled.

(ii) The statement of the physical constraints.

(iii) The specification of the performance criterion.

Given a dynamical model and the corresponding cost functional[6], there are basically two ways for solving the OCP, one is the Pontryagyn MP [24] and the other is the Bellman's PO [25]. In this thesis we are concentrated on the use of Bellman's principle.

DP is the technique derived of the PO, the basic idea for the application of DP is easy to understand for discrete systems, where it can be seen as a discrete, multistage optimization problem in the sense that at each of the finite set of times, a decision is chosen from a finite number of decisions based on certain optimization criterion [7].

In continuous systems, the DP approach leads to the called HJB equation. The PO and the DP approach for continuous systems is presented in the next sub-sections following by the description of the ANN theory for approximation, these three topics are the keystone for the understanding of this thesis.

[6] In some works is also named performance index or performance criterion.

1.3.1 Bellman's Principle of Optimality

In his book, Richard Bellman wrote the following sentence:

In place of determining the optimal sequence of decisions from the fixed state of the system, we wish to determine the optimal decision to be made at any state of the system. Only if we know the latter, do we understand the intrinsic structure of the solution [10].

The principle of Bellman in a simplified way sets that if a path is optimal from a point A to a point C passing through a point B, the path from B to C is also an optimal one. Consider the following example: The supposed optimal path for a multistage decision process is depicted in blue color in Figure 1.1. Now consider that the decision made in point A results in the segment $\overline{AB}$ with the cost J_{AB} and the remaining decisions yield segment $\overline{BC}$.

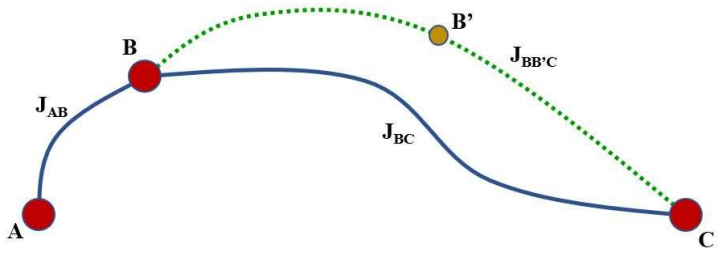

Figure 1.1: Two possible paths from the point A to the point C.

The minimum cost from the point A to the point C is:

$$J^*_{AC} = J_{AB} + J_{BC}.$$

Then, consider the following proposition.

Proposition 1 (PO for multistage decision)**.** *If $\overline{ABC}$ is the optimal path from A to C, then $\overline{BC}$ is the optimal path from B to C.*

The previous statement can be seen as if one can "*break*" the complete optimal path into smaller segments, these segments are optimal and if one finds the optimal costs of the segments, then one can obtain the optimal cost of the complete path.

Proof. By contradiction suppose that the segment $\overline{BB'C}$ is the optimal path from point B to C. Hence, the following inequality is valid:

$$J_{BB'C} < J_{BC},$$

then,

$$J_{AB} + J_{BB'C} < J_{AB} + J_{BC} = J^*_{AC}. \tag{1.1}$$

However, (1.1) can be satisfied only by violating the condition that $\overline{ABC}$ is the optimal path from A to C. □

In other words, the cost functional has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

The keystone theory of the OP and DP was presented in [10] by Bellman. The theory related with its application in control theory can be consulted in [7, 9, 11–13, 22, 26], where in a simplify statement the OP is described as:

An optimal policy has the property that no matter what the previous decision (i.e., controls) have been, the remaining decisions must constitute an optimal policy with regard to the state resulting from those previous decisions.

To describe the OP applied to OCPs for continuous systems, it is important to define formally an OCP. This is a basic theory, however, it is included in this thesis in order to define properly the DP for the robust case in the next chapters.

Optimal Control Problem

Recalling the aforementioned concepts, an OCP needs the mathematical description of the plant or process, the restrictions and the cost functional. The mathematical description for different class of dynamical systems can be represented by the following ODE:

$$\dot{x}(t) = f(t, x, u), \;\; x(t_0) = x_0, \tag{1.2}$$

here $x \in \mathbb{R}^n$ is the state, $u \in U \subset \mathbb{R}^m$ is the control, $t \in \mathbb{R}_+$ is the time, $t_0 \in \mathbb{R}_+$ is the initial time and $x_0 \in \mathbb{R}^n$ is the initial state and the vector field $f\colon \mathbb{R}_+ \times \mathbb{R}^n \times U \to \mathbb{R}^n$ described the dynamics. The conditions for the OCP in this case could be a desired fixed time $t_1 \in \mathbb{R}_+$ and the set U of admissible controls, for this example, the set is restricted to the piecewise continuous signals. The cost functional can be represented in the Bolza[7] form, which is a more general cost functional including the state, input and time.

$$J(t,x,u) = \int_t^{t_1} L(s,x(s),u(s))\,ds + K(x(t_1))\,, \tag{1.3}$$

where $t_1 \in \mathbb{R}_+$ is the fixed final time, $t \in [t_0,t_1)$, $L\colon \mathbb{R}_+ \times \mathbb{R}^n \times U \to \mathbb{R}$ is the running cost[8], and $K\colon \mathbb{R}^n \to \mathbb{R}$ is the terminal cost.

Now that the three main elements to formulate an OCP are described, the formal problem statement is:

To find the input $u^* \in U$ (the OC) such that, this control drives the plant (1.2) *through a state $x^*(t)$ that minimize (or maximize) the cost functional* (1.3)*.*

$$J(t,x,u) \to \min_{u\in U},\ \ \forall t \in [t_0,t_1)\,. \tag{1.4}$$

The PO gives the sufficient conditions to obtain the solution to this OCP using a DP approach (this approach will be described in the following subsection). The next definition of the VF[9] is the key to define the PO for its application in OCPs.

Definition 1 (Value Function)**.** *The function $V(t,x)$ defined for any $t \in [t_0,t_1)$ and $x \in \mathbb{R}^n$ by*

$$V(t,x) := \inf_{u_{[t,t_1)}} J(t,x,u)\,,$$

satisfying the boundary condition

$$V(t_1,x) = K(x(t_1)),$$

is called the VF of the OCP.

[7] There are other class of cost functionals for OCPs, such as Mayer forms, Lagrangian costs and modified Bolza costs with discount terms (exponential functions), saturation functions and Barrier functions [7, 9, 11, 12, 26].

[8] Also known as the Lagrangian [12].

[9] Notice that the OCP is studied for the minimization and the existence of the minimum is not assumed, the infimum is used instead in the definition.

Then the PO can be expressed using the VF, this is the base of DP and to derive the HJB equation.

Proposition 2 (PO for the OCP)**.** *For every* $(t,x) \in [t_0,t_1) \times \mathbb{R}^n$*, the VF satisfies the equation*

$$V(t,x) = \inf_{u_{[t,t_2)}} \left\{ \int_t^{t_2} L(s,x(s),u(s))\,ds + V(t_2,x(t_2)) \right\}, \ \ \forall t_2 \in (t,t_1]. \tag{1.5}$$

Proof. On one hand, by the Definition 1, the following inequality is valid

$$V(t,x) \leq J(t,x,u), \ \ \forall u \in U, \ \ \forall t \in [t_0,t_1), \ \ \forall x \in \mathbb{R}^n.$$

Therefore, substituting (1.3) and considering the boundary condition (terminal cost).

$$V(t,x) \leq \int_t^{t_2} L(s,x(s),u(s))\,ds + J(t_2,x(t_2);u(\cdot)), \ \ \forall t_2 \in (t,t_1].$$

This inequality is valid for any $u \in U$. Taking the infimum over $u(\cdot)$.

$$V(t,x) \leq \bar{V}, \tag{1.6}$$

where $\bar{V}$ is denoting the right hand side of (1.5). On the other hand, for any $\varepsilon > 0$, there exists a control $u_\varepsilon \in U$ such that the derived state x_ε of that control drives

$$V(t,x) + \varepsilon \geq J(t,x,u_\varepsilon), \ \ \forall t \in [t_0,t_1).$$

Then,

$$V(t,x) + \varepsilon \geq \int_t^{t_2} L(s,x_\varepsilon(s),u_\varepsilon(s))\,ds + K(x_\varepsilon(t_2)) \ \ \forall t_2 \in (t,t_1].$$

Therefore, the following inequality is valid

$$V(t,x) + \varepsilon \geq \bar{V}. \tag{1.7}$$

Letting $\varepsilon \to 0$, the inequalities (1.6) and (1.7) imply (1.5). □

In the following subsection it is described the application of the PO to solve an OCP, this drives to the DP approach and the formulation of the HJB equation.

1.3.2 Dynamic Programming

The representation of the PO in an infinitesimal form gives as a result the HJB equation, this is sum up in the following Theorem.

Theorem 1. *Suppose that the VF (Definition 1) is $V \in C^1([t_0,t_1)\times\mathbb{R}^n)$. Then, the VF $V(t,x)$ is a solution to the following terminal value problem of a first order PDE, known as the HJB equation, associated with the OCP* (1.4)*, $\forall t \in [t_0,t_1)$, $\forall x \in \mathbb{R}^n$.*

$$\begin{aligned} -\frac{\partial}{\partial t}V(t,x) + \sup_{u\in U} H\left(-\nabla V(t,x), x(t), u(t), t\right) = 0, \\ V(t_1,x) = K(x(t_1)), \end{aligned} \tag{1.8}$$

here the function $H\colon \mathbb{R}^n\times\mathbb{R}^n\times U\times[t_0,t_1)\to\mathbb{R}$ is known as the Hamiltonian and is defined by

$$H\left(-\nabla V(t,x), x(t), u(t), t\right) := -\nabla^\top V(t,x) f(t,x,u) - L(t,x,u). \tag{1.9}$$

Proof. **Necessary condition.** Considering the definition of the VF and (1.3),

$$V(t,x) = \inf_{u_{[t,t_1)}} \left\{ \int_t^{t_1} L(s,x(s),u(s))\,ds + K\left(x(t_1)\right) \right\}. \tag{1.10}$$

By subdividing the interval of the integral, we obtain

$$V(t,x) = \inf_{u_{[t,t_1)}} \left\{ \int_t^{t+\Delta t} L(s,x(s),u(s))\,ds + \int_{t+\Delta t}^{t_1} L(s,x(s),u(s))\,ds + K\left(x(t_1)\right) \right\}.$$

Then, using the PO,

$$V(t,x) = \inf_{u_{[t,t+\Delta t)}} \left\{ \int_t^{t+\Delta t} L(s,x(s),u(s))\,ds + V(t+\Delta t, x(t+\Delta t)) \right\}.$$

Under the smoothness assumption on the VF, we can express $V(t+\Delta t, x(t+\Delta t))$ using its Taylor series expansion about (t,x), then, we obtain,

$$\begin{aligned} V(t,x) = \inf_{u_{[t,t+\Delta t)}} \Bigg\{ & \int_t^{t+\Delta t} L(s,x(s),u(s))\,ds + V(t,x) \\ & + \frac{\partial}{\partial t}V(t,x)\Delta t + \nabla^\top V(t,x) f(t,x,u)\Delta t + o(\Delta t) \Bigg\}. \end{aligned}$$

Simplifying the previous expression by gathering terms and considering in the infimum the associated terms with the control,

$$0 = \inf_{u_{[t,t+\Delta t)}} \left\{ \int_t^{t+\Delta t} L(s,x(s),u(s))\,ds + \nabla^\top V(t,x) f(t,x,u)\Delta t + o(\Delta t) \right\} + \frac{\partial}{\partial t}V(t,x)\Delta t.$$

Dividing the previous equation by Δt and then $\Delta t \to 0$, we obtain

$$0 = \inf_{u \in U} \left\{ L(t,x,u) + \nabla^\top V(t,x) f(t,x,u) \right\} + \frac{\partial}{\partial t} V(t,x). \tag{1.11}$$

Considering the definition of the Hamiltonian and the definition of supremum and infimum[10], equation (4.32) can be rewritten as (1.8), the boundary final value can be obtained by setting $t = t_1$ in (1.10).

Sufficient condition. Suppose that a C^1 function $\hat{V}\colon [t_0,t_1) \times \mathbb{R}^n \to \mathbb{R}$ satisfies the HJB equation, then,

$$-\frac{\partial}{\partial t}\hat{V}(t,x) = \inf_{u \in U} \left\{ L(t,x,u) + \nabla^\top \hat{V}(t,x) f(t,x,u) \right\}. \tag{1.12}$$

Consider that an optimal pair $\hat{u}\colon [t_0,t_1) \to U$ and $\hat{x}\colon [t_0,t_1) \to \mathbb{R}^n$ satisfy the previous equation, with a given initial condition $\hat{x}(t_0) = x_0$,

$$-\frac{\partial}{\partial t}\hat{V}(t,\hat{x}) = L(t,\hat{x},\hat{u}) + \nabla^\top \hat{V}(t,\hat{x}) f(t,\hat{x},\hat{u}).$$

Using the fact that the sum of the time partial derivative of $\frac{\partial}{\partial t}\hat{V}(t,\hat{x})$ and the inner product $\left\langle \nabla \hat{V}, f \right\rangle$ is indeed the full time derivative, we can rewrite the previous equation as follows

$$L(t,\hat{x},\hat{u}) + \frac{d}{dt}\hat{V}(t,\hat{x}) = 0,$$

by the integration of this equation with respect to the time $t \in [t_0,t_1)$, we obtain

$$\int_{t_0}^{t_1} L(t,\hat{x},\hat{u})\, dt + \hat{V}(t_1,\hat{x}(t_1)) - \hat{V}(t_0,\hat{x}(t_0)) = 0.$$

Rewriting this equality, such that,

$$\hat{V}(t_0,x_0) = \int_{t_0}^{t_1} L(t,\hat{x},\hat{u})\, dt + \hat{V}(t_1,\hat{x}(t_1)),$$

which is equal to,

$$\hat{V}(t_0,x_0) = J(t_0,x_0,\hat{u}).$$

[10]The existence of an optimal control has not been assumed, when it (u^*) exists, the infimum in the previous calculations can be replaced by a minimum, whch is achived with u^*.

Now, if we consider other pair $x \in \mathbb{R}^n$ and $u \in U$ satisfying (1.12), the following inequality is valid

$$-\frac{\partial}{\partial t}\hat{V}(t,x) \leq L(t,x,u) + \nabla^\top \hat{V}(t,x) f(t,x,u).$$

Therefore, using similar procedure as before (gathering the partial derivatives to rewrite in terms of the full time derivative and then integrating), we obtain

$$\hat{V}(t_0,x_0) \leq J(t_0,x_0,u). \tag{1.13}$$

Inequalities (2) and (1.13) show that the input $\hat{u}$ gives the cost $\hat{V}(t_0,x_0)$ and any other control u can not obtain a smaller cost. □

The HJB equation provides the solution for the OCP for general nonlinear systems. However, finding the analytic solution of the HJB equation[11] is a difficult task, moreover if we do not have the complete information of the plant. The most logical approach is to propose the form of the solution and prove if it satisfy the equation. One way to approximate the solution is the use of ANN structures, in the following section, the approximation capabilities of ANN are discussed.

1.3.3 Approximation using ANN

An additional difficulty for solving the HJB equation is added because of the robustness asked to the OC. The proposition is to obtain an approximated solution based on ANNs[12].

The use of ANN structures in diverse control and estimation problems (see for example [27–30]) has presented reliable results. One of the motivations of the use of such structures is because ANNs are recognized by their universal approximation capacity [19]. This characteristic has been used to approximate static maps, time dependent and multivariable nonlinear functions [31, 32]. Moreover, ANNs can be used to approximate the trajectories associated with

[11] Satisfying the smoothness with respect to time and state simultaneously.

[12] The VF must be smooth enough with respect to its both arguments to fulfill sufficient and necessary conditions of optimality. The solution must consider the specific characteristics of the HJB trajectories (positiveness, continuity and differentiability).

the solutions of ODEs and PDEs [33, 34]. Some particular variants of these approximations appeared in the last decade to approximate the solutions of the HJB equations[13].

An ANN can be seen as a simplified mathematical representation of the neural connections in a living being. The unit or the simplest element in an ANN structure is the artificial neuron[14], which can be mathematically represented by the following equation

$$z = \varphi\left(\sum_{j=1}^{p} \omega_j x_j + b\right), \tag{1.14}$$

where $x_j \in \mathbb{R}$ are the input signals, $\omega_j \in \mathbb{R}$ are the synaptic weights, $b \in \mathbb{R}$ is the bias and $\varphi\colon \mathbb{R} \to \mathbb{R}$ is the activation function. Graphically the artificial neuron is depicted in Figure 1.2, where it is compared with a biological cell.

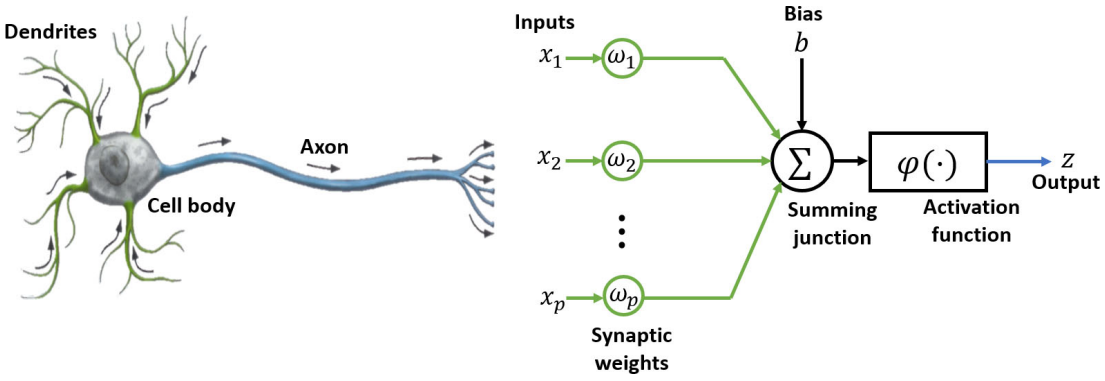

Figure 1.2: Comparison between the information flow on a biological neuron and its nonlinear model.

In a biological neuron, the dendrites collect the electrical signals. This is emulated in the artificial model with a set of synapses, each of which is characterized by a weight. The cell body contains the nucleus and organelles, in the artificial model the body is structured by an adder summing the weighted input signals. In the artificial model the activation function characterizes the output, while in the biological neuron is the axon which passes the electrical signals on to another cell.

[13] The approach is called NDP, this is addressed in the State of art section.

[14] The ANN unit is a basic mathematical representation of a biological neuron.

There exists different types of activation functions: threshold, piecewise-linear, sigmoids, polynomials, Radial-Basis, Gaussian, Wavelet, among others [35]. The selection of the activation functions and the structure depends on the application of the ANN [27]. The structure or architecture is related with the configuration of the neurons, moreover, the configuration is intimately related with the learning algorithm to train the net or the adaptive algorithm to find the synaptic weights.

The common applications of ANNs are: pattern association [36], pattern recognition [37], filtering [38], smoothing [39], prediction [40], control [30], observation [41], system identification [42], beamforming [43] and function approximation [44]. The application of ANNs for function approximation has been studied with different methodologies. There exist a relation between the ANN topology and the studied structures in the general approximation theory.

In the literature of ANN, some works are devoted to the theory of approximation by superposition of nonlinear functions [45, 46], other research use the application of the Kolmogorov theory [19, 47–49] or the application of the Stone–Weierstrass Theorem [50], also the frequency domain approach using Fourier series approximation [51, 52], among other perspectives.

For example, the issue of the ANN structure to achieve a local approximation for nonlinear mappings depends on the basis of activation functions, the number of layers, the configuration or in general the complexity of the net. In [31], the approximation bounds for a class of ANN with sigmoidal activation functions are derived, the structure is a one-layer with n nodes. In [53] the role of the input dimension is considered in the estimation for the bounds of approximation errors, the structure of the ANN has a perceptron and Gaussian radial computational units. The quality of the approximation remains as an interest of research. In [54] the bound of the approximation error for a single hidden layer feed-forward ANN with fixed weights is analyzed.

Based on all the aforementioned properties, ANN could be implemented as feasible approximations for the trajectories corresponding to well-defined solutions of either ODEs or PDEs. Some particular variants of these approximations appeared in the last decade to approximate the solutions of the HJB equations (see e.g. [15, 55–59]). These solutions usually did not satisfy the constraints of the HJB trajectories[15]. Moreover, the approximation capacity of the ANN is

[15] Positiveness, continuity, differentiability and zero vanishing.

only justified if the number of *artificial neurons* in the network tents to infinity.

In this thesis, the activation functions are selected as sigmoid functions. The selection of such function is based on several results [31, 45, 51], where the approximation of nonlinear multivariable functions on a compact set using sigmoidal functions has been proved.

Definition 2. *A sigmoid function is bounded, differentiable and has a non-negative derivative. For this study, the sigmoid function* $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ *is defined as*

$$\varphi(z) := \left(1 + e^{-\left(y^\top z + \theta\right)}\right)^{-1}, \tag{1.15}$$

where, $z \in \mathbb{R}^n$, $y \in \mathbb{R}^n$ *and* $\theta \in \mathbb{R}_+$.

The universal approximation property of ANNs claims that a continuous function can be approximated on a compact set using weighted superposition of nonlinear functions[16]. The following theorem can be recalled to show the approximate realization of nonlinear functions using sigmoidal functions:

Theorem 2. *[45] Let* $\varphi(\cdot)$ *be any continuous sigmoidal function. Then, finite sums of the form*

$$\Gamma(x) = \sum_{j=1}^{N} w_j \varphi\left(y_j^\top x + \theta_j\right), \quad w_j \in \mathbb{R},\ y_j \in \mathbb{R}^n,\ \theta_j \in \mathbb{R}_+, \tag{1.16}$$

are dense in $C([0,1]^n)$. *In other words, given any* $f \in C([0,1]^n)$ *and* $\varepsilon > 0$, *there is a sum,* $\Gamma(x)$ *of the above form , for which*

$$|\Gamma(x) - f(x)| < \varepsilon, \quad \forall x \in [0,1]^n, \tag{1.17}$$

where $C([0,1]^n)$ *is the space of continuous functions on* $[0,1]^n$ *which denotes the* n*-dimensional unit cube.*

Notice that (1.16) is the structure of an ANN with one internal layer and the needed parameters are w_j, y_j and θ_j. Based on such arguments, in this work, the parameters y_j and θ_j are

[16] The nonlinear functions are the activation functions such as, radial basis functions, polynomials, and sigmoidal functions [17].

randomly selected and the adjustment laws for the parameters w_j are proposed[17] in order to approximate the nonlinear function that satisfies the HJB equation.

In the following section, some results related with the problem presented in this thesis, including the works devoted to the approximation of the HJB solution using ANNs, are described.

1.4 State of art

The classification of OCPs is given by the class of system (linear, nonlinear, continuous, discrete, with uncertainties, etc), the time (finite horizon or infinite horizon) and by the selected approach (the Maximum principle [61] and the PO [10]).

An OCP may consider the optimization of the performance index in the infinite horizon [62], usually, these solutions are preferred because the gain of the OC depends on algebraic solutions of matrix equations such as the Riccati one [7]. Nevertheless, if the OCP considers the finite horizon problem, then it is usual that time-varying gains modify the control action [22]. In the DP theory, the OCs are presented as close-loop solutions. This characteristic makes these controllers more useful within the automatic control area [11, 63].

One of the major difficulties in solving OCPs relies on the assumption that the accurate mathematical model of the plan is available, which seems to be unfeasible. Indeed, the application of OC solutions in real systems has been limited because of the lack of accurate enough models, in some cases, the dynamic description of the system that must be controlled is partially known and it is affected by external perturbations. In consequence, the classical OC approaches cannot be applied straightforwardly [13]. The alternative of robust controllers appeared as an option to deal with modelling uncertainties and the effect of bounded perturbations.

Robust controllers are designed to enforce the convergence of a practical equilibrium point. There are several theories to design robust controls such as: sliding mode controllers [64, 65], attractive ellipsoid based controllers [66, 67] and ANNs based controllers [68, 69]. The combination of robust and optimal control theories is not common because they are dealing with systems

[17] In [29] and [60], the selection of y_j and θ_j is based on a uniform distribution using random parameters and the universal approximation property holds.

which have dissimilar characteristics.

Under some given conditions of the uncertain system that should be regulated, a degree of robustness has been considered in the design of OCs. Systems with parametric uncertainties, multimodel representation [22], or bounded perturbations [21] have been controlled sub-optimally. In general, sub-optimal robust approaches tried to compensate the unknown section of the system exactly, but the effect of the uncertainties over the optimization criterion has not been completely researched. Some other robust-optimal control design approaches use the OC estimated by the DP theory, but assuming that the system is not perturbed. The degree of robustness is evaluated using the variation of the performance index by neglecting the uncertain section of the system [70].

Independently of the uncertainties in the model, the application of the DP theory leads to construct the gain control in terms of the solution of the HJB equation. This equation may have an analytic solution if the system has an exact mathematical model of the plant to be controlled and no perturbations affect the dynamics. The HJB equation when uncertainties and disturbances affect the system has analytic solutions in very specific cases. Indeed, just a few works characterized the effect of perturbations/uncertainties on the HJB solution [55, 56, 71]. Even more, the stability (in some well-defined sense) of the equilibrium point enforced by the sub-optimal control has not been clearly explained yet.

One manner to overcome the problem of finding the HJB solution (maybe not exact but close enough to it) uses the approximate DP concept. This technique has been proposed to attain a feasible realization of robust OCs based on a feasible approximation of the HJB solution. There are diverse applications of the approximate DP, including discrete approximations for the domain of the HJB equation.

In [15] and [72] the authors develop a method to approximate the HJB solution in two steps. In the first step, by means of successive approximations, they reduce the equation to a linear PDE and in the second step, they propose a Galerkin's approximation. In [73] and [74] the approximation of the solution is obtained using a discretization technique. In [75, 76] and [77], the HJB equation is reduced to a state dependent Riccati equation, [78] presents a survey of state dependent Riccati equations. The authors in [79, 80] and [81] use the theory of viscosity solutions

to obtain the VF of specific HJB equations. In [82], the authors give an analytical-approximate solution using He's polynomials based on the called homotopy perturbation method which is a combination of the classical perturbation technique and topological concepts (homotopy). Some of the studies working on approximations of the HJB solution do not consider the quality of the approximation on the optimality conditions and in some of them the plant is considered nonlinear but all the elements are assumed as known.

In recent years, diverse OC designs considered approximations for the VF leading to get a controller that enforces the states to track trajectory which corresponds to the minimum value of the performance index. This methodology is called NDP if the approximation uses ANNs [57, 59, 83].

The use of ANNs to find approximated optimal solutions has been also studied in different works. ANNs are mathematical structures with the property of approximation for numerous class of multivariable nonlinear functions [19]. The number of applications of these algorithms is extensive, e.g. ANNs have been used for classification tasks [37], non-parametric and parametric identification [42,84], observers and controllers design [30], approximation of nonlinear mappings (static and continuous) [46], approximation of solutions for different class of equations such as ODEs and PDEs [85].

Based on all the aforementioned properties, the solution for the HJB equation has been approximated using ANNs structures, e.g., in [86], a nearly optimal solution is obtained with the approximated VF of a generalized HJB equation considering saturated inputs in infinite horizon. The authors of [57] prove the convergence of the ANN approximation using Kronecker matrix methods, the adjustment of the parameters is made off-line. In [58] is proposed an on-line method using policy iteration, meanwhile, [59] presents an ANN solution for discrete-time nonlinear control-affine systems. This ANN approximation must apply different learning paradigms: reinforcement learning (RL), adaptive critics, integral RL or reinforcement Q learning which are unsupervised methods [17,87–89]. In particular, NDP has been used to solve finite horizon OCPs in systems with uncertainties, for example the works [18, 55, 71, 90].

The aforementioned works consider an exact model for the plant, however, the presence of uncertainties or perturbations is a common problem in control design. In [22], a robust optimal

control approach considering the Maximum principle, for the case of multimodel systems is studied. In [55, 71, 91] and [92], an approximated solution using ANNs is proposed for systems with uncertainties. The solutions presented usually do not satisfy the constraints of the HJB trajectories. Moreover, the approximation capacity of the ANN is only justified when the number of elements in the network tents to infinity.

In this thesis, we consider a finite horizon OCP for systems with uncertainties or/and bounded perturbations. The approach to obtain the control law is DP and the HJB equation solution is approximated with ANNs. The method to deal with the unknown dynamics is based in [22], where a robust optimal control is obtained for the case of multi-model systems.

1.5 Contributions

The **main contributions** of this work are as follows:

- The formulation of the finite horizon OCP for continuous systems with presence of uncertainties is presented by setting a min-max DP approach. In the cost functional, the uncertainties are not included explicitly and the boundaries for the admissible set of uncertainties are taken into account instead.

- The necessary and sufficient conditions[18] for optimality of the min-max DP approach are given.

- An ANN structure is designed for the approximation of the HJB equation solution derived from the min-max DP approach. The learning laws for the adjustment of the parameters are obtained by the study of the HJB equation under the approximated solution. In addition, this paper presents the analysis including the approximation error for the optimality conditions.

- The implementation of an off-line algorithm to ensure the convergence of the approximation error is described (training process).

[18] Necessary and sufficient conditions considering the VF definition as solution of the HJB equation.

- The stability analysis of the plant affected by uncertainties in closed loop with the designed control law is realized by means of the Lyapunov theory, using the structure of the proposed ANN.

- In addition, the approximation capabilities of ANN are used to design an identifier for homogeneous nonlinear systems. This result is the initial study of the application of ANNs for homogeneous systems and that can be used to apply NDP for a more general type of systems.

Homogeneity is a kind of symmetry under which an object remains consistent with respect to a certain scaling or dilation. Homogeneous systems can be utilized for local approximations [93,94] or set-valued extensions [95,96] of nonlinear control systems. In particular, some models of process control [97], non-holonomic mechanical systems [98] and systems with frictions [95] are homogeneous or at least locally homogeneous.

One of the main features of homogeneous systems is that the local analysis (on the unit sphere, for example) can be extended to the whole state space [99–101].

Based on all the above mentioned homogeneity properties and the wide application of the non-parametric identifiers for controlling uncertain systems, in this thesis, the design of a non-parametric identifier with a homogeneous structure using Differential Neural Networks (DNNs), which are a class of ANNs with a continuous evolution of its states, is presented.

Among the contributions of this result, is presented the extension of the approximation theorem of ANNs for this class of systems. The idea was to extend this result for a future application in the approximation of solutions for other class of equations such as PDEs and of course the HJB equation.

1.6 Structure

The content of this thesis is structured in the following chapters:

Chapter 2 formulates the OCP describing the class of system considered in this thesis and the class of unknown elements in the system. We consider two class of system, the first has

a structure with an additive perturbation or uncertainty and the second considers an additive perturbation or uncertainty and a multiplicative one associated with the input. The problem considering the perturbations or uncertainties is setting with a min-max approach. Then, the PO is used to obtain a min-max DP methodology, deriving the concept of a *Robust* VF and the respective max-min HJB equation for each case. The complete methodology is summarized in the last section of the chapter.

After the formulation of the max-min HJB equation, in **Chapter 3** an ANN approximated solution for each case is presented. A Corollary is used to demonstrate the approximation capabilities of the ANN structure approximating the *robust* VF. Using the proposed approximation and the max-min HJB equation for each case, the learning laws are derived and a numerical off-line algorithm to adjust the free parameters or the initial conditions in the learning laws is presented. In the last section of this chapter the analysis stability for the equilibrium of the systems is made using as a Lyapunov function candidate the approximated *Robust* VF.

After these results, in **Chapter 4**, other application of the ANN approximation capabilities is presented. Here is important to mention that the main goal of this chapter was to use the ANN and the properties of homogeneous systems, such as, most of local properties can be extended globally. This chapter presents an ANN identifier for a class of homogeneous systems, this work is planned to extend the use of ANN along with homogeneous properties for other applications, for example the approximation of solutions for ODE's and PDE's as the HJB equation.

The work of the aforementioned chapter was develop in an research internship during the PhD at the **Institut National de Recherche en Informatique et en Automatique** in Lille France under the supervision of doctors **Andrey Polyakov** and **Denis Efimov**.

The numerical results of the theory presented in the previously mentioned chapters are presented in **Chapter 5**. The chapter is divided in three sections, one for each result. The numerical simulations includes comparisons with other approaches.

The general **conclusions** and remarks of this thesis are presented after Chapter 5 with the topics that are planned to be studied in a recent future.

At the end, in the appendix, the proofs of some results are stated and the list of publications derived of this thesis.

Chapter 2

The min-max Dynamic Programming

In this chapter, the considered class of systems with uncertain models is described and the finite horizon OCP is formulated according to the min-max optimization approach. The main optimality conditions related with the max-min HJB equation are stated. At the final section, the proposed methodology to obtain the optimal arguments for the max-min HJB equation is described. Then, to summarize the proposed approach, an integrated scheme representing all the steps to solve the OCP is presented.

The class of systems considered in this work satisfies:

$$\dot{x} = Ax + Bu + \eta_i z_i, \quad x(t_0) = x_0, \quad t \in [t_0, T], \quad t_0 \in \mathbb{R}_+, \quad i = a, b, \tag{2.1}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in U_{adm} \subset \mathbb{R}^m$ is the control input in the admissible set U_{adm} defined below. The matrices $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$ are constant known matrices (the column rank of matrix B is m, with $m \leq n$).

The term η_i is defined as $\eta_i = [F_i(x,t), G_i(x,t)]$, where the functions $F_i \colon \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^{n \times n}$, $G_a \colon \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n$ and $G_b \colon \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^{n \times m}$ are uncertain but satisfy the following assumption,

Assumption 1. *The admissible set containing the uncertainties of* (2.1) *is characterized as:*

$$\Psi^{adm} = \{\eta_i : \|F_i(t,x)\|_F \leq \alpha_i, \|G_i(t,x)\|_F \leq \beta_i\}, \tag{2.2}$$

$\forall t \in \mathbb{R}_+$, $\forall x \in \mathbb{R}^n$, *with* $0 < \alpha_i < +\infty$ *and* $0 < \beta_i < +\infty$.

The variable z_i is used here to take into account two type of uncertainties affecting the nonlinear system (2.1):

a) the first case when the system is affected by an additive uncertain term, then $z_a = \left[x^\top 1\right]^\top$,
b) the second case when the system is affected by both a multiplicative uncertainty to the input and an additive uncertain term to the dynamics of x, then, $z_b = \left[x^\top u^\top\right]^\top$.

The following assumptions are assumed to be valid throughout this thesis:

Assumption 2. *The admissible control* $u \in U_{adm} \subset \mathbb{R}^m$ *is piece-wise continuous over* $[t_0, T]$.

Assumption 3. *The pair* (A, B) *is controllable in the Kalman sense.*

The class of systems with uncertain models (2.1) could represent different types of nonlinear systems[1] such as mechanical, biological and chemical, among others [4].

2.1 Robust Optimal Control Problem

The main purpose of this work consists in designing the control $u^* \in U_{adm} \subseteq \mathbb{R}^m$ which minimizes a given cost functional J, for (2.1), subjected to the presence of uncertainties characterized by (2.2).

The general form of the cost functional considered in this study obeys the Bolza form [12]:

$$J(t_0, x_0; u(\cdot)) = h_0(x(T)) + \int_{t_0}^{T} h(x(t), u(t))dt. \tag{2.3}$$

[1] Notice that (2.1) is in fact an affine nonlinear form with respect the input general system and the nonlinear unknown elements are presented using the extended linearization [102] or state-dependent coefficient representation [103]. As a result, the system is factored into a linear-like structure with state dependent matrices.

The cost functional depends on the state and control. The form of function $h : \mathbb{R}^n \times U_{adm} \to \mathbb{R}_+$ on (2.3) satisfies the following quadratic structure:

$$h(x,u) = x^\top Qx + u^\top Ru, \tag{2.4}$$

where $Q \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$ are positive definite and symmetric matrices. The continuous function $h_0 : \mathbb{R}^n \to \mathbb{R}$ defines the terminal condition.

In a classical OCP, if the plant to be controlled has a complete known model then, the OC solution consists in finding the explicit stabilizing u^* which optimizes the given functional. However, in this work, the model (2.1) is affected with some uncertain nonlinear functions which prevents finding the eact calculus of u^*. Hence, one of the contributions of this study is the design of a robust optimal solution for the worst case associated to the class of admissible uncertainties, then a min-max approach. Hence, the proposed controller considers the maximization over the admissible set of the uncertainties. To solve this, an alternative approach to characterize the problem of designing the OC is given in terms of the following min-max problem [22]:

$$\max_{\eta_i \in \Psi^{adm}} J(t_0, x_0; u(\cdot)) \to \min_{u(\cdot)\in U_{adm}[t_0,T]},$$
$$\text{subject to (2.1).}$$

According to the Dynamic Programming Approach, the presented finite-horizon OCP entails in finding the solution for the following PDE (known as the Hamilton-Jacobi-Bellman equation),

$$-\frac{\partial V(t,x)}{\partial t} + \max_{u\in U_{adm}} \min_{\eta_i \in \Psi^{adm}} H\left(-\nabla V(t,x), x, u, \eta_i\right) = 0. \tag{2.5}$$

The function $H : \mathbb{R}^n \times \mathbb{R}^n \times U_{adm} \times \Psi^{adm} \to \mathbb{R}$ is the Hamiltonian, which was defined in (1.9), for the class of system in this OCP, that is

$$H\left(-\nabla V(t,x), x, u, \eta_i\right) = -\nabla^\top V(t,x)\left(Ax + Bu + \eta_i z_i\right) - h\left(x,u\right). \tag{2.6}$$

The function $V : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ is the VF and $t \in [t_0, T]$. In view of the uncertainties effect on the dynamics of (2.1), in this study, the following definition of *robust* VF is considered:

Definition 3. *The robust VF is defined for any* $(t,x) \in [t_0,T) \times \mathbb{R}^n$ *by:*

$$V(t,x) = \min_{u\in U_{adm}} \max_{\eta_i \in \Psi^{adm}} J(t,x;u(\cdot)), \tag{2.7}$$

with the boundary condition $V(T,x) = h_0(x)$.

The max-min equation (2.5) is derived from the PO. This robust version of OC design is useful for the case of multi-model systems (finite dimensional case) or for systems with uncertainties such as (2.1) (infinite dimensional case). The following sub-section presents the main theorem which proves how to derive the max-min estimation on the Hamiltonian considering the *robust* VF associated to the proposed cost functional. Such result implies that the proposed *robust* VF is the solution of a modified (robustified) HJB equation.

2.2 Max-min HJB Equation

In the previous chapter, we studied that the DP method [13] provides sufficient conditions[2] on the optimality for an admissible control with the application of the HJB equation. This study takes advantages of such property to derive the proposed robust OC. The following Theorem derives the corresponding HJB equation for systems with bounded perturbations or parametric uncertainties such as (2.1).

Theorem 3. *Suppose that* $V(t,x)$ *in* (2.7)*, is continuously differentiable with respect to its two arguments, then, it is a solution of* (2.5)*,* $(t,x) \in [t_0,T) \times \mathbb{R}^n$*, with the boundary condition* $V(T,x) = h_0(x(T))$*.*

Proof. Consider a control $u \in U_{adm}$ and the following representation based on the PO (see e.g. [10]) for the robust VF with $s \in [t_0,T)$:

$$V(s,y) = \min_{u \in U_{adm}} \max_{\eta_i \in \Psi^{adm}} \left\{ \int_s^{\hat{s}} h(x(t),u(t))dt + V(\hat{s},x(\hat{s})) \right\},$$
$$\forall \hat{s} \in [s,T], \quad x(s) = y.$$

Therefore, by the definition of the minimum operator, the following inequality is valid:

$$V(s,y) \leq \max_{\eta_i \in \Psi^{adm}} \left\{ \int_s^{\hat{s}} h(x(t),u(t))dt + V(\hat{s},x(\hat{s})) \right\}, \ \forall \hat{s} \in [s,T].$$

Using the fact that $\min\{a\} = -\max\{-a\}$, one can obtain:

$$V(s,y) + \min_{\eta_i \in \Psi^{adm}} \left\{ -\int_s^{\hat{s}} h(x(t),u(t))dt - V(\hat{s},x(\hat{s})) \right\} \leq 0, \ \forall \hat{s} \in [s,T].$$

[2] DP states necessary conditions assuming smoothness properties on the VF.

Multiplying by $\frac{1}{\hat{s}-s}$ and considering the Definition 3, the following inequality holds:

$$\frac{1}{\hat{s}-s}\left[\min_{\eta_i\in\Psi^{adm}}\left\{V(s,y)-V\left(\hat{s},x(\hat{s})\right)-\int_s^{\hat{s}}h(x(t),u(t))dt\right\}\right]\leq 0. \tag{2.8}$$

Using the Mean Value Theorem in (2.8), we obtain:

$$-\frac{\partial V(t,x)}{\partial t}+\min_{\eta_i\in\Psi^{adm}}\left\{-\nabla^{\top}V(t,x)\left(Ax+Bu+\eta_i z_i\right)-h(x(t),u(t))\right\}\leq 0,\ \forall\hat{s}\in[s,T],$$

which implies:

$$0\geq-\frac{\partial V(t,x)}{\partial t}+\max_{u\in U_{adm}}\min_{\eta_i\in\Psi^{adm}}\left\{H\left(-\nabla V(t,x),x,u,\eta_i\right)\right\}. \tag{2.9}$$

On the other hand, for any $\varepsilon>0$ and s close to $\hat{s}$, there exists a control signal $u(\cdot):=u_{\varepsilon,\hat{s}}(\cdot)\in U_{adm[s,T]}$ for which:

$$V(s,y)+\varepsilon(\hat{s}-s)\geq\max_{\eta_i\in\Psi^{adm}}\left\{\int_{t=s}^{\hat{s}}h(x(t)_{\varepsilon,\hat{s}},u(t)_{\varepsilon,\hat{s}})dt+V(\hat{s},x(\hat{s}))\right\}. \tag{2.10}$$

Subtracting $V(s,y)$ from both sides of (2.10) and multiplying by $\frac{1}{\hat{s}-s}$, one can obtain:

$$\varepsilon\geq\frac{1}{\hat{s}-s}\left[\max_{\eta_i\in\Psi^{adm}}\left\{\int_{t=s}^{\hat{s}}h(x(t)_{\varepsilon,\hat{s}},u(t)_{\varepsilon,\hat{s}})dt+V(\hat{s},x(\hat{s}))-V(s,y)\right\}\right]. \tag{2.11}$$

Using the Fundamental Calculus Theorem on (2.11) and multiplying it by -1 at both sides

$$-\varepsilon\leq\frac{1}{\hat{s}-s}\left[-\max_{\eta_i\in\Psi^{adm}}\left\{\int_s^{\hat{s}}-H\left(-\nabla V(t,x),x_{\varepsilon,\hat{s}},u_{\varepsilon,\hat{s}},\eta_i\right)dt\right\}-\int_{t=s}^{\hat{s}}\frac{\partial V(t,x)}{\partial t}dt\right]. \tag{2.12}$$

Then, considering the definition of the minimum and maximum, the inequality (2.12) can be rewritten as follows

$$-\varepsilon\leq\frac{1}{\hat{s}-s}\int_s^{\hat{s}}\left[\min_{\eta_i\in\Psi^{adm}}\left\{H\left(-\nabla V(t,x),x_{\varepsilon,\hat{s}},u_{\varepsilon,\hat{s}},\eta_i\right)\right\}-\frac{\partial V(t,x)}{\partial t}\right]dt.$$

Therefore, taking the maximum on the right-hand side, the following inequality is also valid

$$-\varepsilon\leq\frac{1}{\hat{s}-s}\int_s^{\hat{s}}\left[\max_{u\in U_{adm}}\min_{\eta_i\in\Psi^{adm}}\left\{H\left(-\nabla V(t,x),x_{\varepsilon,\hat{s}},u_{\varepsilon,\hat{s}},\eta_i\right)\right\}-\frac{\partial V(t,x)}{\partial t}\right]dt. \tag{2.13}$$

Let $s\to\hat{s}$ in (2.13), hence, we obtain:

$$-\varepsilon\leq\max_{u\in U_{adm}}\min_{\eta_i\in\Psi^{adm}}\left\{H\left(-\nabla V(t,x),x,u,\eta_i\right)\right\}-\frac{\partial V(t,x)}{\partial t}. \tag{2.14}$$

Considering (2.9) and (2.14) and taking $\varepsilon\to 0$ the proof is completed. □

The previous proof uses the condition on the robust VF $V\in C^1$. Notice also that the construction of the previous proof states a kind of necessary conditions. The following section states the sufficient conditions of the optimality.

2.3 Verification Rule

Consider the optimal admissible value for the uncertainties $\eta_i^* \in \Psi^{adm}$, such that the following optimization is solved,

$$H\left(-\nabla V(x,t),x,u,\eta_i\right) \rightarrow \min_{\eta_i \in \Psi^{\text{adm}}},$$

and the control law given by

$$u^*(\cdot) := u^*\left(\eta_i^*, x, \nabla V(t,x)\right), \tag{2.15}$$

be a robust optimal solution, such that,

$$H\left(-\nabla V(x,t),x,u,\eta_i^*\right) \rightarrow \max_{u \in U_{\text{adm}}}, \tag{2.16}$$

and suppose that we can obtain a VF solution $V^* \in C^1$ to the HJB equation (2.5), then

$$-\frac{\partial}{\partial t}V^*(t,x) + H\left(-\nabla V^*(t,x),x,u,\eta_i\right) = 0, \quad V^*(T,x) = h_0(x(T)), \tag{2.17}$$

which for any $(t,x) \in [t_0,T) \times \mathbb{R}^n$ is unique and smooth.

Considering the OC (u^*) in the system (2.1), there exists a solution $x^* \in \mathbb{R}^n$ satisfying the ODE (2.1).

Suppose the pair (u^*,x^*), $u^*_{[t_0,T)} \in U_{adm}$, $x^*_{[t_0,T)} \in R^n$ satisfies (2.17), then the definition of the Hamiltonian (2.6) for this OCP implies $V^*(t,x^*(t))$ satisfies the following ODE

$$\frac{d}{dt}V^*(t,x^*(t)) = -h\left(x^*,u^*\right).$$

Integrating this equality by $t \in [t_0,T]$ leads to the relation

$$V^*(T,x^*(T)) - V^*(t_0,x^*(t_0)) = -\int_{t_0}^{T} h\left(x^*(t),u^*(t)\right)dt. \tag{2.18}$$

The last equation could be rewritten by considering the final condition for VF as:

$$V^*(t_0,x_0) = h_0(x^*(T)) + \int_{t_0}^{T} h\left(x^*(t),u^*(t)\right)dt. \tag{2.19}$$

By Definition 3, the equation (2.19) means exactly that (x^*,u^*,η_i^*) is an optimal triplet and u^* is the OC.

The max-min HJB to obtain a robust-like OC needs the optimal arguments η_i^* and u^* and then, the *robust* is the VF solution. The following section describes the proposed methodology to obtain the optimal values η_i^* and u^*.

2.4 Optimal arguments for the max-min HJB

The first step to solve (2.5) consists in finding η_i^* such that the Hamiltonian is minimized.

2.4.1 Optimal value for the uncertain element (η_i^*)

One can consider only the terms related to η_i due to the structure of (2.5). Then, the solution for the first step can be achieved by solving the following equivalent optimization problem:

$$\eta_i^* = \underset{\eta_i \in \Psi^{adm}}{\arg\min} \left\{ -\nabla^\top V(t,x) \eta_i z_i \right\}.$$

The solution for this problem was gotten with two different approaches, one for each case of z_i.

a) **Additive uncertain term.**

For the case of z_a, consider the following partial optimization problem to characterize η_a^*:

$$\eta_a^* = \underset{\eta_a \in \Psi^{adm}}{\arg\min} \left\{ -\nabla^\top V(t,x) \eta_a z_a \right\}. \tag{2.20}$$

Using the inner product, the equation (2.20) is equivalent to

$$\eta_a^* = \underset{\eta_a \in \Psi^{adm}}{\arg\min} \left\{ -\langle \eta_a^\top \nabla V(t,x), z_a \rangle \right\}. \tag{2.21}$$

The application of the Cauchy-Schwartz inequality on the inner product in (2.21) leads to

$$-\langle \eta_a^\top \nabla V(t,x), z_a \rangle \geq -\left\| \eta_a^\top \nabla V(t,x), \right\| \|z_a\|,$$

then, using the triangle inequality[3] on the previous inequality, such that

$$-\langle \eta_a^\top \nabla V(t,x), z_a \rangle \geq -\left\| \eta_a^\top \right\|_F \|\nabla V(t,x)\| \|z_a\|,$$

and considering the Assumption 1 regarding the admissible set of uncertainties Ψ^{adm} (2.2), one gets

$$-\langle \eta_a^\top \nabla V(t,x), z_a \rangle \geq -\sqrt{\phi} \|\nabla V(t,x)\| \|z_a\|, \tag{2.22}$$

[3] Considering the adequate consistent norm in the space of matrices [104,105], the Frobenious norm is compatible with the Euclidean vector norm.

where $\phi = \alpha_a^2 + \beta_a^2$, with α_a and β_a given in (2.2).

The inequality (2.22) can be used to obtain the exact value of η_a^* taking into account the equality:

$$\eta_a^* = \sqrt{\phi}\nabla V(t,x) z_a^\top \left\| \nabla V(t,x) z_a^\top \right\|^{-1}.$$

b) **Additive uncertain term and multiplicative uncertainty with the input.**

For the case of z_b, consider the following partial optimization problem to characterize η_b^*:

$$\eta_b^* = \underset{\eta_b \in \Psi^{adm}}{\arg\min} \left\{ -\nabla^\top V(t,x)\eta_b z_b \right\}.$$

Using the Λ-inequality[4], the terms related to η_b in the Hamiltonian can be bounded as follows:

$$-\nabla^\top V(t,x)\eta_b z_b \geq -0.5\left(\|\nabla V(t,x)\|_\Lambda^2 + \|\eta_b z_b\|_{\Lambda^{-1}}^2 \right) .$$

Considering that $\Lambda = I_n$, and the compatibility of norms in metric spaces [105], the following inequality is also valid:

$$-\nabla^\top V(t,x)\eta_b z_b \geq -0.5\left(\|\nabla V(t,x)\|^2 + \|\eta_b\|_F^2 \|z_b\|^2 \right).$$

Considering (2.2), one can obtain

$$-\nabla^\top V(t,x)\eta_b z_b \geq -0.5\|\nabla V(t,x)\|^2 - \gamma\|z_b\|^2,$$

where $\gamma = 0.5\left(\alpha_b^2 + \beta_b^2\right)$. Then, the optimal argument is

$$\eta_b^* = \nabla V(t,x) z_b^\top \left(0.5\left\|z_b\right\|^{-2} + \gamma \left\|\nabla V(t,x)\right\|^{-2} \right).$$

The second step consists on finding the OC u_i^*, which is obtained by solving the following equation

$$u_i^* = \underset{u \in U_{adm}}{\arg\max} \left\{ \min_{\eta_i \in \Psi^{adm}} H\left(-\nabla V(t,x), x, u, \eta_i\right) \right\}.$$

The following subsection describes the methodology to obtain the OC.

[4]For all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ the inequality $2x^\top y \leq \|x\|_\Lambda^2 + \|y\|_{\Lambda^{-1}}^2$ holds with any $\Lambda \in \mathbb{R}^{n\times n}$, $\Lambda = \Lambda^\top$, positive definite [105].

2.4.2 Optimal control value (u^*)

The methodological difference to obtain the optimal values η_i^* is mainly because the second case considers a multiplicative unknown element with respect to the input. The optimal value η_b^* has separated (additive) elements $(\nabla V(t,x), z_b)$ while in the first approach, for the case of η_a^*, these elements are multiplicative. Then, after the substitution of the optimal value η_i^* in each case, we use the following approach to obtain the OC u_i^*.

a) **Additive uncertain term.**

In this case, the OC is obtained directly [5] because the unknown elements are not related with the input, then

$$u_a^* = -0.5R^{-1}B^\top \nabla V(t,x). \tag{2.23}$$

b) **Additive uncertain term and multiplicative uncertainty with the input.**

By the substitution of η_b^*, equation (2.5), for this case, is equivalent to:

$$\begin{aligned} -\frac{\partial V(t,x)}{\partial t} + \max_{u\in U_{adm}} \Big\{ & -\nabla^\top V(t,x)Ax - \nabla^\top V(t,x)Bu_b - 0.5\,\|\nabla V(t,x)\|^2 \\ & -\gamma\|z_b\|^2 - \|x\|_Q^2 - \|u_b\|_R^2 \Big\} = 0. \end{aligned} \tag{2.24}$$

Considering the terms related to the control (u) in (2.24), the maximization problem over the Hamiltonian is equal to:

$$u_b^* = \arg\max_{u_b\in U_{adm}} \left\{ -\nabla^\top V(t,x)Bu_b - \gamma\|z_b\|^2 - \|u_b\|_R^2 \right\}.$$

Considering the control u on vector z_b and gathering terms,

$$u^* = \arg\max_{u\in U_{adm}} \left\{ -\nabla^\top V(t,x)Bu - \|u\|_\Phi^2 \right\},$$

where $\Phi = \gamma I_m + R$. Therefore, the optimal control u^* is given by:

$$u^* = -0.5\Phi^{-1}B^\top \nabla V(t,x). \tag{2.25}$$

The HBJ equation for the cases, considering the optimal arguments is given by:

[5] By direct differentiation respect to u.

a) **Additive uncertain term.**

Substituting both u^* and η^* on (2.5), such that

$$-\frac{\partial V(t,x)}{\partial t} - \nabla^\top V(t,x)Ax + \frac{1}{4}\nabla^\top V(t,x)BR^{-1}B^\top \nabla V(t,x) \\ -\|x\|_Q^2 - \sqrt{\phi\left(\|x\|^2+1\right)}\,\|\nabla V(t,x)\| = 0. \tag{2.26}$$

b) **Additive uncertain term and multiplicative uncertainty with the input.**

The substitution of (2.25) in (2.24) yields to

$$-\frac{\partial V(t,x)}{\partial t} - \nabla^\top V(t,x)Ax - \|\nabla V(t,x)\|_{\Phi_2}^2 - \|x\|_{\Phi_3}^2 = 0, \tag{2.27}$$

where $\Phi_2 = 0.5I_n - 0.25B\Phi^{-1}B^\top$ and $\Phi_3 = Q + \gamma I_n$.

The min-max DP approach to obtain a robut-like OC solution is summarized in the following section.

2.5 Procedure summary of the min-max DP approach

Table 2.1 summarizes all the steps of the proposed methodology to solve the min-max DP approach of the analyzed OCP.

Table 2.1: Procedure summary

Step	Description	
1	Problem Formulation	Given the plant as (2.1), the performance index (2.3) and the initial condition $x(t_0) = x_0$, find the OC u^*, considering the unknown elements η_i characterized by (2.2). $$\max_{\eta_i \in \Psi^{adm}} J(t_0, x_0; u(\cdot)) \rightarrow \min_{u(\cdot)\in U_{adm}[t_0,T]},$$ subject to (2.1) and (2.2).

2	PO for the min-max problem, the min-max DP	After the formulation problem, setting the optimal conditions for the DP approach using the PO for the min-max problem. The *robust* VF is defined for any $(t,x) \in [t_0,T) \times \mathbb{R}^n$ by $$V(t,x) = \min_{u \in U_{adm}} \max_{\eta_i \in \Psi^{adm}} J(t,x;u(\cdot)).$$
3	The max-min HJB	The use of the VF definition to obtain the PDE (max-min HJB equation) $$-\frac{\partial V(t,x)}{\partial t} + \max_{u \in U_{adm}} \min_{\eta_i \in \Psi^{adm}} H\left(-\nabla V(t,x), x, u, \eta_i\right) = 0.$$
4	Minimization of the Hamiltonian	To find the optimal argument η_i^* for the PDE of the previous step, minimize respect to η_i. $$\eta_a^* = \sqrt{\phi} \nabla V(t,x) z_a^\top \left\| \nabla V(t,x) z_a^\top \right\|^{-1},$$ and $$\eta_b^* = \nabla V(t,x) z_b^\top \left(0.5 \left\| z_b \right\|^{-2} + \gamma \left\| \nabla V(t,x) \right\|^{-2} \right).$$
5	Maximization of the Hamiltonian	With a fixed η_i^*, to find the OC u^* for the PDE of the step 3, maximize respect to u. For a) $$u^* = -0.5 R^{-1} B^\top \nabla V(t,x).$$ For b) $$u^* = -0.5 \Phi^{-1} B^\top \nabla V(t,x).$$

6	The optimal HJB equation	To obtain the HJB equation after the substitution of the optimal arguments of the steps 4 and 5. $$-\frac{\partial V(t,x)}{\partial t} + H\left(-\nabla V(t,x), x, u^*, \eta_i^*\right) = 0.$$
7	Solution of the HJB equation (VF)	To solve the PDE of the step 6 with terminal condition $V(T,x(T)) = h_0(x(T))$, obtaining the respective VF.
8	Computation of the OC	To compute the gradient of the solution (*robust* VF) obtained in the previous step, i.e., $\nabla V(t,x)$.

Notice that the structures of the Step 5 (OC for (2.1)) need the gradient (Step 8) of the solution for the max-min HJB in Step 7. An additional difficulty for solving the max-min HJB is added because of the robustness asked to the control solution. The Steps 7 y 8 are described in the following Chapter, that is, the ANN approximation for the solution of this max-min HJB equation and the computation of the gradient.

Chapter 3

ANN approximation for the max-min HJB

This chapter describes the ANN approximation for the max-min HJB equation solution. The ANN approximation theorem (universal approximation property) is used to demonstrate the approximation capabilities of the ANN structure approximating the *robust* VF. The description of such an approximation for both classes of systems, a) the case of additive uncertainties and b) the case of additive and input multiplicative uncertainties is given. In addition, the stability analysis for the controlled systems using the approximated VF as a Lyapunov function candidate is also presented.

This part of the thesis analyses two major aspects regarding the ANN application as a feasible approximation for the *Robust* VF ($V(t,x)$): the first deals with the justification of the requested properties to derive the approximate solution of the max-min HJB and the second studies the approximation error enforced by the ANN to the *robustified* version of the VF.

The following section describes the proposed ANN structure, which approximates the solution for each case of the two aforementioned cases for the max-min HJB equation.

3.1 ANN approximated VF

The VF $V(t,x)$ for linear systems without perturbations or uncertainties is a quadratic form of the state vector x. Then, the problem of solving (2.5) modifies to ensure the existence of an uniform positive definite solution for a differential Riccati equation [12, 63, 106]. Inspired on this solution, the proposed approximation considers a similar quadratic form of the state complemented with a suitable ANN topology.

Consider the following representation of (2.26) and (2.27):

$$V(t,x) = V_a(t,x) + \tilde{V}(t,x), \tag{3.1}$$

where $V_a : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ is the ANN-based approximation and $\tilde{V} : \mathbb{R}_+ \times \mathbb{R}^n$ is the approximation error. Therefore, the approximation structure is proposed as

$$V_a(t,x) = V_{NN}(t,x) + x^\top P(t)x, \tag{3.2}$$

with $P : \mathbb{R}_+ \to \mathbb{R}^{n\times n}$ being positive definite $\forall t \in [t_0, T)$ and uniformly bounded with respect to t, let $x \in \mathbb{R}^n$ are the states of (2.1) and V_{NN} is the selected ANN structure. The following Corollary justifies the approximation property (using the ANN structure) for the solution of (2.26) and (2.27).

Corollary 1. *Consider that the Robust VF* (2.7) *satisfies the conditions of Theorem 3 and consider $P : \mathbb{R}_+ \to \mathbb{R}^{n\times n}$ to be a positive definite matrix $\forall t \in \mathbb{R}_+$. Then, for any $\varphi_a \in \mathbb{R}_+$, there exists an integer p defining the number of neurons of the V_{NN} structure and a set of p output weights, such that:*

$$|\tilde{V}(t,x)| \leq \varphi_a, \quad \forall x \in \mathbb{R}^n, \quad \forall t \in \mathbb{R}_+.$$

Proof. The absolute value of the approximation error can be represented as:

$$|\tilde{V}(t,x)| = |V_{NN}(t,x) - v_1(t,x)|, \quad \forall t \in \mathbb{R}_+, \quad \forall x \in \mathbb{R}^n,$$

where $v_1(t,x) = V(t,x) - x^\top P(t)x$. The function $v_1(t,x)$ is continuous for all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}_+$. Applying the Theorem 2, the proof is completed. □

The approximate function V_{NN} satisfies a single-layer ANN topology. In this study, the learning laws design is made using the min-max HJB equation for each class of systems (additive and multiplicative uncertainties).

3.1.1 ANN solution for the case a)

Assume that $V_{NN}(t,x)$ represents an approximate solution associated to the max-min HJB (2.26) given by the following ANN structure

$$V_{NN}(x,t) = \omega(t)\sigma(x). \tag{3.3}$$

The structure of $\omega \in \mathbb{R}^+$ and $\sigma \in \mathbb{R}^+$ is proposed as follows:

$$\omega(t) = \tilde{\omega}^\top(t)\tilde{\omega}(t), \quad \sigma(x) = \tilde{\sigma}^\top(x)\tilde{\sigma}(x).$$

The function $\tilde{\omega} : \mathbb{R}_+ \to \mathbb{R}^p$ corresponds to the weights vector and $\tilde{\sigma} : \mathbb{R}^n \to \mathbb{R}^p$ is the ANN activation vector based on sigmoid activation functions. This particular structure ensures the positiveness of the approximate *Robust* VF. For this case, the sigmoid function satisfies the Definition 2. Therefore, the max-min HJB equation (2.26) based on the approximated function $V_a(t,x)$ considering $\tilde{V}(t,x) \to 0$, is governed by

$$\begin{aligned} &-x^\top \dot{P}(t)x - \dot{\omega}(t)\sigma(x) - \omega(t)\nabla^\top\sigma(x)Ax - 2x^\top P(t)Ax \\ &+\frac{1}{4}\left(\omega(t)\nabla^\top\sigma(x) + 2x^\top P(t)\right)BR^{-1}B^\top\left(\omega(t)\nabla\sigma(x) + 2P(t)x\right) \\ &-\sqrt{\phi\left(\|x\|^2+1\right)}\,\|\omega(t)\nabla\sigma(x) + 2P(t)x\| - \|x\|_Q^2 = 0. \end{aligned} \tag{3.4}$$

The equation (3.4) can be separated in time dependent differential equations using the rewriting of quadratic terms $x^\top \mathrm{Ric}(P)x$ as follows,

$$\mathrm{Ric}(P) = -\dot{P}(t) - P(t)A + A^\top P(t) + P(t)BR^{-1}B^\top P(t) - Q - \frac{4\sqrt{\phi\left(\|x\|^2+1\right)}}{z(\omega(t),P)}P^2(t), \tag{3.5}$$

and the elements $\tilde{\omega}^\top\tilde{W}$, where

$$\begin{aligned} \tilde{W} = &-\frac{d\tilde{\omega}(t)}{dt} + \frac{\tilde{\omega}(t)\nabla^\top\sigma(x)}{2\sigma(x)}\left[-Ax + \frac{1}{4}\omega(t)BR^{-1}B^\top\nabla\sigma(x)\right. \\ &\left.+BR^{-1}B^\top 2P(t)x - \frac{4\sqrt{\phi\left(\|x\|^2+1\right)}}{z(\omega,P)}P(t)x - \frac{\omega(t)\sqrt{\phi\left(\|x\|^2+1\right)}}{z(\omega,P)}\cdot\nabla\sigma(x)\right], \end{aligned} \tag{3.6}$$

Here $z(\omega(t),P) = \|\omega(t)\nabla\sigma(x) + 2P(t)x\|$, considering $\tilde{W} = 0$ and $\mathrm{Ric}(P) = 0$. The simultaneous on-line numerical solution of (3.5) and (3.6), and the numerical solution for the initial conditions define an approximated numerical solution of the max-min HJB.

3.1.2 ANN solution for the case b)

Let consider the following structure for the approximation of (2.27):

$$V_{NN}(t,x) = \sum_{i=1}^{p} \omega_i(t)\sigma_i(x), \tag{3.7}$$

where $\omega_i : \mathbb{R}_+ \to \mathbb{R}_+$ are defined as $\omega_i(t) = (\tilde{\omega}_i(t))^2$, $\tilde{\omega}_i : \mathbb{R}_+ \to \mathbb{R}$ are the adjustment weights of the ANN structure. Therefore, the HJB equation using (3.7) has the approximated solution governed by

$$\begin{aligned} &-x^\top \dot{P}(t)x - \sum_{i=1}^{p} \dot{\omega}_i(t)\sigma_i(x) - 2x^\top P(t)Ax - \sum_{i=1}^{p} \omega_i(t)\nabla^\top\sigma_i(x)Ax - x^\top\Phi_3 x \\ &-4\left[x^\top P(t) + \sum_{i=1}^{p} \omega_i(t)\nabla^\top\sigma_i(x)\right]\Phi_2\left[P(t)x + \sum_{j=1}^{p} \omega_j(t)\nabla\sigma_j(x)\right] = 0. \end{aligned} \tag{3.8}$$

The equation (3.8) can be rewritten gathering the quadratic terms with respect the state and the elements associated to the weights

$$\begin{aligned} &-x^\top\left[\frac{P}{dt}(t) + P(t)A + A^\top P(t) + 4P(t)\Phi_2 P(t) + \Phi_3\right]x \\ &-\sum_{i=1}^{p} \tilde{\omega}_i(t)\left[2\frac{d}{dt}\tilde{\omega}_i(t)\sigma(x) + \tilde{\omega}_i(t)\nabla^\top\sigma_i(x)Ax + 8\tilde{\omega}_i(t)\nabla^\top\sigma_i(x)\Phi_2 P(t)x\right. \\ &\left. + 4\tilde{\omega}_i(t)\nabla^\top\sigma_i(x)\Phi_2\sum_{j=1}^{p}\left(\tilde{\omega}_j(t)\right)^2\nabla\sigma_j(x)\right] = 0. \end{aligned} \tag{3.9}$$

Then, (3.9) can be separated in the following two differential equations:

$$\frac{P}{dt}(t) = -P(t)A - A^\top P(t) - P(t)4\Phi_2 P(t) - \Phi_3, \tag{3.10}$$

and

$$\frac{d}{dt}\tilde{\omega}_i(t) = -\frac{\tilde{\omega}_i(t)\nabla^\top\sigma_i(x)}{2\sigma_i(x)}\left[Ax + 8\Phi_2 P(t)x + 4\Phi_2\sum_{j=1}^{p}\left(\tilde{\omega}_j(t)\right)^2\nabla\sigma_j(x)\right]. \tag{3.11}$$

Equations (3.5) and (3.6) are the learning laws associated with the ANN solution for (2.26) and equations (3.10)-(3.11) are the learning laws associated with the ANN approximation of (2.27). From these learning laws, we obtain the structure of $V_a(t,x)$ for both classes of systems with uncertain models. Notice that the gradient of this approximated *Robust* VF ($\nabla V_a(t,x)$) can be used for realizing the corresponding OCs.

The initial condition and the methodology for the implementation of the OCs are described in the following section.

3.2 Numerical adjustment of the free ANN parameters

The solution for the class of *robust* OC requires the implementation of a special class of recurrent algorithm (Algorithm 1) to adjust the initial conditions of the weights participating in the ANN structure. The realization of this method is equivalent to the regular training process of ANN-based approximation solutions.

The algorithm implements a routine to check the terminal condition of the approximated *Robust* VF. If the expected value is not gotten, then an adjust for the parameters of the learning laws is enforced by the proposed algorithm.

The numerical method used a mixed strategy realizing the evolution of the states regulated by the sub-optimal controller. Each sequence was evaluated by adjusting only the initial condition of the weights adjustment laws[1].

Figure 3.1 depicts the scheme to represent the interaction between the numerical solution and the OC implementation.

[1] Algorithm 1 implements a class of Levenberg-Marquardt method [107] for inverse problems (parameter identification). In [108], the authors present the proofs for the global convergence and they do not assume that there is a solution with zero residuals, which is a natural condition if the number of activation functions in the ANN is finite.

Algorithm 1 Recurrent algorithm to adjust the initial parameters of the learning laws equations

Start

Variables initialization

κ, μ; Γ; w_a; μ_i; γ_a; i;

$\mu_i > \mu$; $w_i \leftarrow w_a$; $V_i \leftarrow \mu_i$; $\gamma_i \leftarrow \gamma_a$; $\bar{\mu} = \mu_i$;

while $V_i \geq \mu$ **do**

  Exert simulation of $\left(x_i, u_i = -0.5R^{-1}B_0^{\top}\frac{\partial}{\partial x}V(w_i)\right)$;

  $aux_{1i} = \left(\frac{\partial}{\partial w}V(w_i)\right)^{\top}\left(\frac{\partial}{\partial w}V(w_i)\right) + \gamma_i I_p$;

  $aux_{2i} = \left(\frac{\partial}{\partial w}V(w_i)\right)^{\top} V(w_i)$;

  $s_{i+1} = -aux_{1i}^{-1} aux_{2i}$;

  $num_i = V_i^2 - V_i^2(w_i + s_i)$;

  $den_i = V_i^2 - \left(V_i + \frac{\partial}{\partial w}V(w_i)s_i\right)^2 - \gamma_i \|s_i\|^2$;

  $\rho_i = den_i^{-1} num_i$;

  if $\rho_i \geq \kappa$ **then**

    $w_{i+1} = w_i + s_i$;

    $\mu_{i+1} \in \left[\max\left\{\mu, \frac{\bar{\mu}}{\Gamma}\right\}, \bar{\mu}\right]$;

    $\bar{\mu} = \mu_{i+1}$;

  else

    $w_{i+1} = w_i$;

    $\mu_{i+1} = \Gamma\mu_i$;

  end if

  $V_{i+1} = V(w_i)$;

  $\gamma_{i+1} = \mu_{i+1}\left\|\frac{\partial}{\partial w_i}V_{i+1}\right\|^2$;

  $i = i + 1$;

end while

  $u^* = u_i$;

Stop

The vector $w \in \mathbb{R}^p$ is defined as $w = \left[\tilde{\omega}_1(0), \ldots, \tilde{\omega}_p(0)\right]^\top$. Here, the parameter i denotes the number of iteration and it is initialized with $i = 0$. The initial parameters are selected such as, $\kappa \in (0,1)$, $\mu \in \mathbb{R}_+$, $\Gamma > 1$ $w_a \in \mathbb{R}^p_+$, $\mu_i > \mu$, $\gamma_a \in \mathbb{R}_+$. The following section presents the Lyapunov stability analysis using the approximated *Robust* VF.

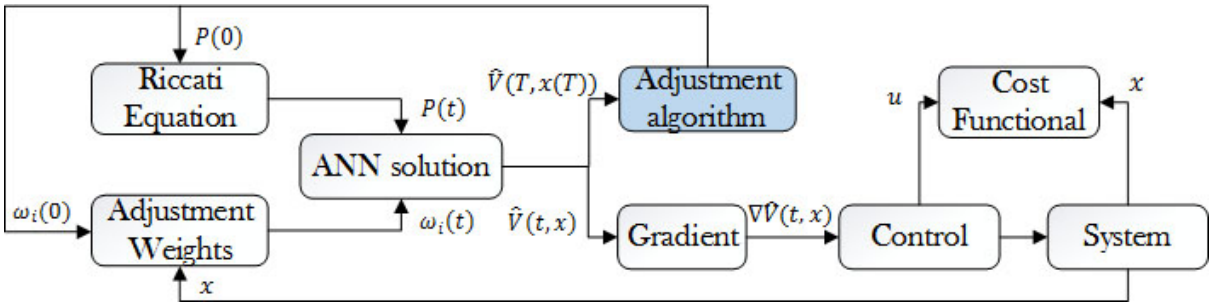

Figure 3.1: Structure of the routines for the numerical simulation and the adjustment of the parameters.

3.3 Practical stability analysis using the OC solution

The presence of uncertainties F_i and G_i has an impact on the value estimation for the performance index (2.3) evaluated over the trajectories of the nonlinear system (2.1). Therefore, a robustness analysis for the state trajectories is requested. The stability analysis for the equilibrium point of (2.1) is based on the Lyapunov stability method. In the following sub-sections, the stability analysis for each case is presented.

3.3.1 Lyapunov analysis for the case a)

Proposition 3. *Consider the nonlinear uncertain system given in* (2.1) *with the sub-optimal controller design introduced in* (2.25)*. Suppose that Assumption 1 stands, then, if there is a bounded positive and symmetric matrix P ($P^- \leq P(t) \leq P^+$ with $P^- \in \mathbb{R}^{n\times n}$ and $P^+ \in \mathbb{R}^{n\times n}$ positive defi-*

nite matrices) of the time varying Riccati matrix equation $(\Lambda_1 > 0, \Lambda_1 \in \mathbb{R}^{n\times n})$*:*

$$\begin{aligned} &\dot{P}(t) - \left(P(t)A + A^\top P(t)\right) \\ &+2P(t)BR^{-1}B^\top P(t) - P(t)\Lambda_1 P(t) - Q_2 = Q_4(t), \end{aligned} \tag{3.12}$$

with $Q_4 : \mathbb{R}_+ \to \mathbb{R}^{n\times n}$ *a time-dependent positive definite matrix, and the weights in the ANN approximation* (3.3) *governed by* (3.6) *are bounded by* $\omega \leq \omega^+ \ \forall t \geq 0$ *with* $\sigma(x) \neq 0, \forall x \neq 0$ *then, the origin defines an uniformly practical stable equilibrium point of* (2.1) *[109] with an ultimate bound given by*

$$\beta = \frac{\varepsilon_0}{\sup_{t\geq 0}\lambda_{max}\{Q_4(t)\}}. \tag{3.13}$$

Proposition 1 states sufficient conditions only. The boundedness of the stable trajectory under the robust optimal closed loop control depends on the existence of the positive solution for (3.5). The proposed method can provide just sufficient arguments in view of the main result, which comes from a Lyapunov-like stability analysis.

Proof. Introduce the approximate value function $V_a : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^+$ as a feasible Lyapunov function candidate for (2.1). Based on the assumptions, the candidate energetic function satisfies

$$x^\top P^- x \leq V_a(t,x) \leq x^\top P^+ x + \omega^+ \sigma(x).$$

Notice that both $x^\top P^- x$ and $\omega^+\sigma(x) + x^\top P^+ x$ are both class-$\mathcal{K}$ functions [3]. The full-time derivative of V_a along the trajectories of x and ω satisfies:

$$\dot{V}_a(t,x(t)) = 2x^\top(t)P(t)\dot{x}(t) + x^\top(t)\dot{P}(t)x(t) + \frac{d}{dt}\omega(t)\sigma(x) + \omega(t)\nabla^\top\sigma(x)\dot{x}(t). \tag{3.14}$$

Reorganizing the elements on (3.14), such differential equation yields

$$\dot{V}_a(t,x(t)) = \left(2x^\top(t)P(t) + \omega(t)\nabla^\top\sigma(x)\right)\dot{x}(t) + x^\top(t)\dot{P}(t)x(t) + 2\bar{\omega}^\top(t)\frac{d}{dt}\bar{\omega}(t)\sigma(x). \tag{3.15}$$

The substitution of (2.1) in (3.15) leads to

$$\begin{aligned} \dot{V}_a(t,x(t)) &= x^\top(t)\dot{P}(t)x(t) + 2\bar{\omega}^\top(t)\frac{d}{dt}\bar{\omega}(t)\sigma(x) \\ &+ \left(2x^\top(t)P(t) + \omega(t)\nabla^\top\sigma(x)\right)\left[Ax(t) + Bu(t) + \eta_a z_a\right]. \end{aligned} \tag{3.16}$$

The application of the sub-optimal admissible controller u^* (2.25) transforms (3.16) to

$$\begin{aligned}\dot{V}_a(t,x(t)) = x^\top(t)\left(P(t)A + A^\top P(t)\right)x(t) + 2x^\top(t)P(t)B\left(-\frac{1}{2}R^{-1}B^\top\nabla V(t,x)\right)\\ +2x^\top P(t)\eta_a z_a + \omega(t)\nabla^\top\sigma(x)\left[Ax + B\left(-\frac{1}{2}R^{-1}B^\top\nabla V(t,x)\right)\right]\\ +\omega(t)\nabla^\top\sigma(x)\eta_a z_a + x^\top\dot{P}(t)x(t) + 2\bar{\omega}^\top(t)\frac{d}{dt}\bar{\omega}(t)\sigma(x).\end{aligned} \tag{3.17}$$

The gradient of V can be estimated using the definition of activation functions. Then, the reorganization of the terms in (3.17) is equivalent to:

$$\begin{aligned}\dot{V}_a(t,x(t)) = x^\top\left(P(t)A + A^\top P(t) - 2P(t)BR^{-1}B^\top P(t)\right)x - 2x^\top P(t)BR^{-1}B^\top\omega(t)\nabla\sigma(x)\\ +2x^\top P(t)\eta_a z_a + \omega(t)\nabla^\top\sigma(x)Ax - \frac{1}{2}\omega(t)\nabla^\top\sigma(x)BR^{-1}B^\top\omega(t)\nabla\sigma(x)\\ +\omega(t)\nabla^\top\sigma(x)\eta_a z_a + x^\top\dot{P}(t)x(t) + 2\bar{\omega}^\top(t)\frac{d}{dt}\bar{\omega}(t)\sigma(x).\end{aligned}$$

The application of the Young's inequality [66], justifies the following inequality, straightforwardly from the previous equation

$$\begin{aligned}\dot{V}_a(t,x(t)) \leq x^\top\left(P(t)A + A^\top P(t) - 2P(t)BR^{-1}B^\top P(t)\right)x + \omega(t)\nabla^\top\sigma(x)Ax + x^\top\dot{P}(t)x(t)\\ -\frac{1}{2}\omega(t)\nabla^\top\sigma(x)BR^{-1}B^\top\omega(t)\nabla\sigma(x) - 2x^\top P(t)BR^{-1}B^\top\omega(t)\nabla\sigma(x) + 2\bar{\omega}^\top(t)\frac{d}{dt}\bar{\omega}(t)\sigma(x)\\ +\left(\omega(t)\nabla^\top\sigma(x)\right)\Lambda_2\left(\omega(t)\nabla^\top\sigma(x)\right)^\top + x^\top P(t)\Lambda_1 P(t)x + z_a^\top\eta_a^\top\Lambda_1^{-1}\eta_a z_a + z_a^\top\eta_a^\top\Lambda_2^{-1}\eta_a z_a.\end{aligned}$$

Factorizing with respect to ω and η_a ($\Lambda^{-1} = \Lambda_1^{-1} + \Lambda_2^{-1}$),

$$\begin{aligned}\dot{V}_a(t,x) \leq x^\top\left(\dot{P}(t) + P(t)A + A^\top P(t) - 2P(t)BR^{-1}B^\top P(t) + P(t)\Lambda_1 P(t)\right)x\\ -\omega(t)\left(2x^\top P(t)BR^{-1}B^\top - x^\top A^\top + \frac{1}{2}\nabla^\top\sigma(x)BR^{-1}B^\top\omega(t) - \omega(t)\nabla^\top\sigma(x)\Lambda_2\right)\nabla\sigma(x)\\ +z_a^\top\eta_a^\top\Lambda^{-1}\eta_a z_a + 2\bar{\omega}^\top(t)\frac{d}{dt}\bar{\omega}(t)\sigma(x).\end{aligned}$$

Considering that $F_a^\top\Lambda^{-1}F_a \leq Q_0$ y $G_a^\top\Lambda^{-1}G_a \leq \varepsilon_0$, one gets

$$\begin{aligned}\dot{V}_a(t,x) \leq x^\top\left(\dot{P}(t) + P(t)A + A^\top P(t) - 2P(t)BR^{-1}B^\top P(t) + P(t)\Lambda_1 P(t) + Q_0\right)x\\ -\omega(t)\left(2x^\top P(t)BR^{-1}B^\top - x^\top A^\top + \frac{1}{2}\nabla^\top\sigma(x)BR^{-1}B^\top\omega(t) - \omega(t)\nabla^\top\sigma(x)\Lambda_2\right)\nabla\sigma(x)\\ +2\bar{\omega}^\top(t)\frac{d}{dt}\bar{\omega}(t)\sigma(x) + \varepsilon_0.\end{aligned}$$

Substituting the adjustment laws for the weights, the previous inequality yields to

$$\begin{aligned}\dot{V}_a(t,x) \leq x^\top \left(\dot{P}(t)+P(t)A+A^\top P(t)-2P(t)BR^{-1}B^\top P(t)+P(t)\Lambda_1 P(t)+Q_0\right)x\\ -\omega(t)\left(2x^\top P(t)BR^{-1}B^\top - x^\top A^\top + \frac{1}{2}\nabla^\top\sigma(x)BR^{-1}B^\top\omega(t)-\omega(t)\nabla^\top\sigma(x)\Lambda_2\right)\nabla\sigma(x)\\ +2\bar{\omega}^\top(t)\frac{\tilde{\omega}(t)\nabla^\top\sigma(x)}{2\sigma(x)}\left(-Ax+\frac{1}{4}\omega(t)BR^{-1}B^\top\nabla\sigma(x)+BR^{-1}B^\top 2P(t)x\right.\\ \left.-\frac{\omega(t)\sqrt{\phi\left(\|x\|^2+1\right)}}{z(\omega,P)}\nabla\sigma(x)\right)\sigma(x)+\varepsilon_0-\frac{4\sqrt{\phi\left(\|x\|^2+1\right)}}{z(\omega,P)}P(x)x\sigma(x).\end{aligned}$$

Simplifying the previous expression, one gets

$$\begin{aligned}\dot{V}_a(t,x) \leq x^\top \left(\dot{P}(t)+P(t)A+A^\top P(t)-2P(t)BR^{-1}B^\top P(t)+P(t)\Lambda_1 P(t)+Q_0\right)x\\ -\omega(t)\left[\frac{1}{4}\nabla^\top\sigma(x)B_0R^{-1}B_0^\top\omega(t)-\omega(t)\nabla^\top\sigma(x)\Lambda_2\right]\nabla\sigma(x)+\varepsilon_0\\ -\omega(t)x^\top P(t)\frac{4\sqrt{\phi\left(\|x\|^2+1\right)}}{z(\omega,P)}\nabla\sigma(x)-\nabla^\top\sigma(x)\frac{\omega^2(t)\sqrt{\phi\left(\|x\|^2+1\right)}}{z(\omega,P)}\nabla\sigma(x).\end{aligned}$$

Considering the extended vector $\varphi := \begin{bmatrix} x^\top & \nabla^\top\sigma(x)\end{bmatrix}^\top$ yields to the simplified form given by:

$$\begin{aligned}\dot{V}_a(t,x) \leq -x^\top\left(-\dot{P}(t)-P(t)A-A^\top P(t)+2P(t)BR^{-1}B^\top P(t)-P(t)\Lambda_1 P(t)-Q_2\right)x\\ -\omega(t)\varphi^\top Q_3\varphi+\varepsilon_0,\end{aligned} \tag{3.18}$$

where:

$$\begin{aligned}Q_2 := Q_0+\omega(t)Q_1,\ Q_3(t) := \begin{bmatrix}\Pi_{11} & \Pi_{12}\\ \Pi_{12}^\top & \Pi_{22}\end{bmatrix},\\ \Pi_{12} := 2\frac{\sqrt{\phi\left(\|x\|^2+1\right)}}{z(\omega,P)}P(t),\ \Pi_{11} := Q_1,\ Q_1\in\mathbb{R}^{n\times n}, Q_1=Q_1^\top, Q_1>0\\ \Pi_{22} := \left[\frac{1}{4}B_0R^{-1}B_0^\top-\Lambda_2+\frac{\omega(t)\sqrt{\phi\left(\|x\|^2+1\right)}}{z(\omega,P)}\right]\omega(t).\end{aligned}$$

Considering (3.12), the inequality (3.18) is equivalent to:

$$\frac{d}{dt}V_a(t,x) \leq -x^\top Q_4(t)x-\omega(t)\varphi^\top Q_3\varphi+\varepsilon_0,$$

If one considers the subspace of the state variables $x\in\left\{x^\top Q_4(t)x\geq\varepsilon_0\right\}$ ($Q_4(t)$ is given in (3.12)), then

$$\frac{d}{dt}V_a(t,x)\leq 0$$

In consequence, the states x are ultimately bounded with the bound given in (3.13). □

3.3.2 Lyapunov analysis for the case b)

For the second considered case, let use the following assumption for system (2.1).

Assumption 4. *The uncertainty $G_b(x,t)$ satisfies the following representation:*

$$G_b(x,t) = B\Xi(x,t), \tag{3.19}$$

where $\Xi : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^{m\times m}$ is bounded $\|\Xi(x,t)\|_F \leq \nu$, $\nu > 0$, $\forall t \in \mathbb{R}_+$, $x \in \mathbb{R}^n$.

Theorem 4. *Consider the nonlinear system given in* (2.1) *with the suboptimal control design*

$$u^* = -0.5\Phi^{-1}B^\top \nabla V_a(t,x). \tag{3.20}$$

Suppose that assumptions 1-3 and Assumption4 hold, then, if there exists a bounded positive definite and symmetric matrix P ($P^- \leq P(t) \leq P^+$ with $P^- \in \mathbb{R}^{n\times n}$ and $P^+ \in \mathbb{R}^{n\times n}$ positive definite matrices) solution of (3.10)*, and the weights in the ANN approximation* (3.7) *governed by* (3.11) *are bounded by $|\tilde{\omega}_i| \leq \omega_i^+$ for all $t \in \mathbb{R}_+$ with $\sigma_i(x) \neq 0$ for all $x \neq 0$, and matrix R in* (2.4) *satisfies*

$$R \leq \phi I_m, \tag{3.21}$$

where $\phi = \dfrac{\gamma}{2}\left(\nu^{-2} - \nu^{-1}\sqrt{\nu^{-2}-4} - 2\right)$, with $\gamma = 0.5(\alpha_b^2 + \beta_b^2)$, α_b and β_b are given in Assumption 1 and ν is given in Assumption 4. Then, the origin is an uniformly practical stable equilibrium point of (2.1)*.*

Proof. Consider (3.2) as the Lyapunov function candidate for (2.1). Based on the conditions of Theorem 4, the approximation for the *robust* VF (3.2) satisfies the following sequence of inequalities

$$W_1(x) \leq V_a(t,x) \leq W_2(x), \quad \forall t \in \mathbb{R}_+, \quad \forall x \in \mathbb{R}^n,$$

where $W_1 : \mathbb{R}^n \to \mathbb{R}_+$ is given by $W_1 = \lambda_{\min}\{P^-\}\|x\|^2$ and $W_2 : \mathbb{R}^n \to \mathbb{R}_+$ is given by $W_2 = \lambda_{\max}\{P^+\}\|x\|^2 + \max\limits_{i=1,\ldots,p}\left\{\left(\omega_i^+\right)^2\right\}\sum\limits_{i=1}^{p}\sigma_i(x)$.

The derivative of (3.2) along the trajectories of (2.1) is defined by

$$\dot{V}_a(t,x) = \frac{\partial V_a(t,x)}{\partial t} + \nabla^\top V_a(t,x)Ax + \nabla^\top V_a(t,x)F_b(t,x)x + \nabla^\top V_a(t,x)Bu + \nabla^\top V_a(t,x)G_a(t,x)u.$$

Considering that (3.2) satisfies (2.27) and by the substitution of (3.20) in the previous equation, the following relation holds

$$\dot{V}_a(t,x) = -\|\nabla V_a\|^2_{\Phi_2} - \|x\|^2_{\Phi_3} - \frac{1}{2}\nabla^\top V_a(t,x) B\Phi^{-1}B^\top \nabla V_a(t,x) \\ +\nabla^\top V_a(t,x) F_b(t,x)x - \tfrac{1}{2}\nabla^\top V_a(t,x) G_b(t,x)\Phi^{-1}B^\top \nabla V_a(t,x).$$

Using the Λ-inequality [105] on the element $\nabla^\top V_a(t,x)F_b(t,x)x$, such that,

$$\nabla^\top V_a(t,x)F_b(t,x)x \leq \frac{1}{2}x^\top F_b^\top(t,x)F_b(t,x)x + \frac{1}{2}\nabla^\top V_a(t,x)V_a(t,x).$$

Gathering terms, the following inequality is valid

$$\dot{V}_a(t,x) \leq -\nabla^\top V_a(t,x)\Pi_a \nabla V_a(t,x) - x^\top \Pi_b x - \tfrac{1}{2}\nabla^\top V_a(t,x) G_b(t,x)\Phi^{-1}B^\top \nabla V_a(t,x), \quad (3.22)$$

where $\Pi_a = \frac{1}{4}B\Phi^{-1}B^\top$ and $\Pi_b = Q + \frac{1}{2}\beta_b^2 I_n$.

Applying the same procedure (Λ-inequality) to the third element,
$\frac{1}{2}\nabla^\top V_a(t,x) G_b(t,x)\Phi^{-1}B^\top \nabla V_a(t,x)$, of (3.22), one can obtain

$$\dot{V}_a(t,x) \leq -\nabla^\top V_a(t,x)\Pi_a \nabla V_a(t,x) - x^\top \Pi_b x + \tfrac{1}{4}\nabla^\top V_a(t,x) G_b(t,x)\Lambda G_b^\top(t,x)\nabla V_a(t,x) \\ +\tfrac{1}{4}\nabla^\top V_a(t,x) B\Phi^{-1}\Lambda^{-1}\Phi^{-1}B^\top \nabla V_a(t,x).$$

Selecting $\Lambda = \gamma^{-1} I_m$ in the previous inequality and considering the Assumption 4, one gets

$$V_a(t,x) \leq -\nabla^\top V_a(t,x)\Pi_c \nabla V_a(t,x) - x^\top \Pi_b x, \quad (3.23)$$

with $\Pi_c = \dfrac{1}{4}B\left(\Phi^{-1}R\Phi^{-1} - \nu^2\gamma^{-1}I_m\right)B^\top$. Therefore, considering (3.21), the equilibrium point of (2.1) is uniformly asymptotically stable (see e.g [3], Theorem 4.9). □

The numerical results for both cases: a) additive uncertain term and b) additive uncertain term and multiplicative uncertainty with the input are presented in Chapter 5. The following Chapter explores the ANN approximation capabilities for a class of systems with uncertain models but satisfying a homogeneous property. Such assumption provides a natural way to extend the applicability of the suggested sub-optimal controller.

The presented ANN identifier is devoted to a class of homogeneous system (standard homogeneity). This research was developed during the PhD internship at the Institut National de

Recherche en Informatique et en Automatique (INRIA) in Lille France under the direction of doctors Andrey Polyakov and Denis Efimov. The original idea was to study the ANN approximation capabilities along with the properties of homogeneous structures to approximate solutions for PDEs, which is a goal of this thesis. As a first application of ANN with homogeneous structures, the identifier of the next chapter was obtained. It is planned to continue the research for the application of ANN with homogeneous structures for other class of applications such as NDP in near future.

Chapter 4

ANN Identification for homogeneous systems

This chapter presents the design of a non-parametric identifier for homogeneous systems based on a class of ANNs with continuous dynamics. The study extends the universal approximation property of ANNs for homogeneous systems. The adjustment laws for the weights are obtained from a Lyapunov stability analysis and the ultimate boundedness of the origin for the identification error is formally demonstrated. In addition, under the persistent excitation condition, the boundedness for the error of the weights in the ANN structure design is also proven.

Analysis, design and optimization of dynamic control systems need a valid mathematical model of the plant. However, most of the exiting control systems have model uncertainties. Therefore, the tools looking for a valid mathematical description of the dynamic system are widely demanded and applied.

Nonlinear system identification is a field of control theory [110], which develops algorithms of mathematical modeling of control systems based on input and output signals measured online or/and during some experiments. Identification problem has been tackled using different approaches due to a large class of system models and its inherent complexity. The identifi-

cation process in general needs to handle the available input and output data in order to postulate a model and validate it somehow. Some of the most popular identification techniques are: functional series methods, frequency domain approaches, fuzzy models and ANNs (see e.g. [27, 111–113] for more details).

The class of neural networks with continuous dynamics (DNN) are utilized to get the approximation of dynamical systems [29, 30, 114], since they can be trained on-line (in a real time). In addition, DNNs can process many inputs and outputs, so they are applicable to multi-variable systems. The DNN identification admits the selection of different activation functions, which represent a certain basis for the model of the system in an admissible space, for example, sigmoidal functions, polynomials or radial basis functions [45, 115, 116]. The adjustment of the time-varying parameters (weights) in the DNN according to its structure and the set of activation functions should be adjusted, for example, by a stability analysis based on Lyapunov procedure, see e.g. [28, 117]. In this chapter, the DNN identification algorithms are developed for a specific class of systems: homogeneous nonlinear systems affine in the input.

Homogeneity is a symmetry-like property under which an object remains consistent with respect to a certain scaling or dilation. Homogeneous systems can be utilized for local approximations [93, 94] or set-valued extensions [95, 96] of nonlinear control systems. In particular, some models of process control [97], nonholonomic mechanical systems [98] and systems with frictions [95] are homogeneous or at least locally homogeneous.

Identification problem of homogeneous systems is not well studied in the literature. One of the main features of homogeneous system is that an analysis of its behavior in a whole state space can be reduced to a similar analysis on a unit sphere [99–101]. This feature implies a specific structure of the DNN identifier. The activation functions are selected to approximate the systems on the unit sphere, next due to homogeneity the system model can be expanded to the whole space. The rest of the chapter develops the identifier for a class of homogeneous systems with uncertain dynamics.

The considered nonlinear system with uncertain model is given by the following ordinary

differential equation:

$$\dot{x} = f(x,u), \quad t \in \mathbb{R}_+, \quad x(0) = x_0, \quad f(x,u) := f_0(x) + \sum_{i=1}^{m} f_i(x)u_i, \tag{4.1}$$

where $x(t) \in \mathbb{R}^n$ is the state vector of the system and $x_0 \in \mathbb{R}^n$ is the initial condition.

The function $f_0 \colon \mathbb{R}^n \to \mathbb{R}^n$ is unknown, the input associated functions $f_i \colon \mathbb{R}^n \to \mathbb{R}^n$, $i = 0,1,...,m$ are unknown nonlinear vector fields, $u(t) = [u_1(t),..,u_m(t)]^\top \in \mathbb{R}^m$ is the control input, $m \leq n$. The identifier design for (4.1) is studied under the following basic assumptions:

Assumption 5. *The vector fields f_i, $i = 0,1,...,m$ are continuous on the unit sphere:*

$$S = \{x \in \mathbb{R}^n : \|x\| = 1\}. \tag{4.2}$$

Assumption 6. *The vector fields f_i are homogeneous in the standard sense*[1] *with known homogeneity degrees $\nu_i \in \mathbb{R}$, i.e., $f_i(\lambda x) = \lambda^{\nu_i} f_i(x)$, $\forall x \in \mathbb{R}^n$, $\forall \lambda > 0$, where $i = 0,1,...,m$.*

The standard homogeneity for the given function f means that such a function f is symmetric with respect to dilation $x \mapsto \lambda x$ of its argument[2].

Assumption 7. *The whole state vector of* (4.1) *assumed to be on-line measured, bounded and sufficiently excited by control inputs (the details are given Theorems 5 and 6).*

Assumption 8. *The control inputs are known essentially bounded functions such that*

$$|u_i(t)| \leq U < +\infty, \quad \forall t \in \mathbb{R}_+, \quad i = 1,\ldots,m.$$

[1] The generalized concepts of homogeneity have been developed for other types of dilation $D \colon \mathbb{R} \to \mathbb{R}^{n\times n}$, $x \mapsto D(\lambda)x$, and for both finite and infinite dimensional systems, see e.g. [101, 118–120].

[2] In this thesis, the standard homogeneity property in the finite dimensional case is considered, since any generalized homogeneous system is topologically equivalent to a standard homogeneous one [121, 122].

4.1 System Identification Problem

The identification problem consists on finding an approximate model for the vector field f and its parameters such that, the error between the states of (4.1) and the states of the approximate structure is small (in a certain norm). Therefore, the first step is to represent (4.1) as a valid approximate model taking into account the homogeneity property of f. To solve this first step, we propose a homogeneous DNN structure. The second step consists in designing adaptive laws for the adjustment of weights of this DNN identifier, such that, the identification error

$$e := x - \hat{x} \tag{4.3}$$

between the system states x and the DNN identifier states $\hat{x}$ satisfies $\limsup_{t\to\infty} \|e(t)\| \leq \rho(\varepsilon^+) < +\infty$, where $\varepsilon^+ \in \mathbb{R}_+$ characterizes the best possible approximation of the unknown mapping f by means of the DNN identifier and $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ is a class-$\mathcal{K}$ function. Obviously, if $\varepsilon^+ = 0$ then, the error $e(t)$ must tend to zero[3] as $t \to +\infty$.

4.2 ANNs property for homogeneous systems

The approximation capabilities of ANNs with sigmoidad functions were stated in Theorem 2, which is formulated under the assumption that the ANN has two layers and it has a static structure. The proof of this theorem has its fundamentals on the Stone–Weierstrass and the Kolmogorov approximation theorems [50]. This property is used in DNN structures for the representation of dynamic systems.

Based on the necessity of a well-justified approximation theory, the next result establishes the universal approximation property of the homogeneous ANN structures for the case of systems as the proposed in (4.1).

Corollary 2. *Consider that* (4.1) *satisfies the assumptions 5 and 6. Then, for any* $\varepsilon_i \in \mathbb{R}_+$ *and for any Hurwitz matrix* $A \in \mathbb{R}^{n\times n}$ *there exist* $N_i \in \mathbb{R}$ *and* $W_i^* \in \mathbb{R}^{n\times N_i}$, $i = 0,1,..,m$ *such that:*

$$\|p(x,u)\| \leq \varepsilon_0 \|x\|^{\nu_0} + \sum_{i=1}^{m} \varepsilon_i \|x\|^{\nu_i} |u_i|, \quad \forall x \in \mathbb{R}^n, \forall u \in \mathbb{R}^m, \tag{4.4}$$

[3] $\varepsilon^+ = 0$, i.e., the DNN model may exactly approximate the original system.

where $p(x,u) := f(x,u) - F(x,u)$ *and the approximate function* $F(x,u)$ *is given by*

$$F(x,u) = \|x\|^{\nu_0}\left[\frac{Ax}{\|x\|} + W_0^*\sigma_0\left(\frac{x}{\|x\|}\right)\right] + \sum_{i=1}^{m}\|x\|^{\nu_i}W_i^*\sigma_i\left(\frac{x}{\|x\|}\right)u_i. \tag{4.5}$$

The elements of the vector functions $\sigma_i : \mathbb{R}^n \to \mathbb{R}_i^N$ *are proposed as* (1.15)*.*

Proof. Based on the properties given in Assumption 6, the system (4.1) can be rewritten as follows:

$$\dot{x} = \|x\|^{\nu_0} f_0\left(\frac{x}{\|x\|}\right) + \sum_{i=1}^{m}\|x\|^{\nu_i} f_i\left(\frac{x}{\|x\|}\right)u_i.$$

Notice that, the right-hand side of (4.1) is uniquely identified by its values on the unit sphere (4.2). Consider the functions[4] $\tilde{f}_i : \mathbb{R}^n \to \mathbb{R}^n$, $i = 0,1,...,m$, i.e., $\tilde{f}_i(x) = f_i\left(\frac{x}{\|x\|}\right)$.

The modeling error function $\tilde{f}_i$ is continuous on $\mathbb{R}^n \setminus \{0\}$ due to continuity of f_i on the unit sphere [5]. Applying the Theorem 2 to each component of the vector $\tilde{f}_i$, $i = 1,2,\ldots,m$ and to $\tilde{f}_0(x) - Ax/\|x\|$ on (4.2), and based on the continuity arguments, it can be observed that $W_i^*\sigma_i(x/\|x\|)$ approximates $\tilde{f}_i(x) = f_i(x/\|x\|)$ with an arbitrary small error ε_i. Notice that the application of sigmoidal functions justifies that the approximation error can be made arbitrary small with the proper selection of parameters (scalars and vectors) W_i^*, y_{ij} and θ_{ij}. Taking into account the homogeneity of nonlinear functions $f_i, i = 0,...,m$, the desired global estimate (4.4) holds. □

We assume that the parameters y_{ij} and θ_{ij}, are selected randomly with a uniform distribution, and it is needed to find matrices W_i^* in order to complete the identification of the uncertain homogeneous model (See Definition 2).

Remark 1. *The activation functions* (1.15) *are bounded. Hence, the vectors of the activation functions used for the identification on the unit sphere are bounded in the following sense:*

$$\left\|\sigma_i\left(\frac{x}{\|x\|}\right)\right\| \leq \sqrt{N_i},\ \forall x \in \mathbb{R}^n. \tag{4.6}$$

[4] Associated to the components of the control u, except $\tilde{f}_0$.

[5] The vector field $f_i : \mathbb{R}^n \to \mathbb{R}^n$ satisfying Assumption 6, is locally Lipschitz continuous on $\mathbb{R}^n \setminus \{0\}$ if and only if it satisfies the Lipschitz condition on (4.2), see e.g. [122] and [123]. Similarly, (4.1) satisfying assumptions 1 and 2 has the continuous right-hand side (on the first argument) in $\mathbb{R}^n \setminus \{0\}$. For $\nu_i = 0$, the function f_i may be discontinuous at the origin. Therefore, the developed identifier has to be able to deal with discontinuous models.

Usually, a function f can be approximated by NN structure only on a compact set (see Theorem 2). The homogeneous ANN structure (4.5) gives a global approximation of f (see the formula (4.4)).

4.3 Identification of affine homogeneous control systems

In this section, the identifier design for two cases on the identification problem of (4.1) is presented. The first case considers that the vector fields associated with the input are known and the second case assumes that all vector fields are unknown.

4.3.1 The case of known control gains

Let consider the case when the nonlinear maps $f_i\colon \mathbb{R}^n \to \mathbb{R}^n$, $i = 1,\ldots,m$ associated with the inputs are known and we need to identify only the vector field f_0.

Remark 2. *In Corollary 2, the bound for the approximation error for the complete unknown vector field case is presented. Hence, the approximation error considering f_0 unknown and the control gains $f_i, i = 0,\ldots,m$ as known, leads to the following inequality:*

$$\left\| d\left(\frac{x}{\|x\|}\right)\right\| \leq \varepsilon_0 \ , \ \forall x \in \mathbb{R}^n, \tag{4.7}$$

where $d\left(\frac{x}{\|x\|}\right) = f_0\left(\frac{x}{\|x\|}\right) - A\frac{x}{\|x\|} - W_0^*\sigma_0\left(\frac{x}{\|x\|}\right)$.

Notice that

$$W_0^*\sigma_0\left(\frac{x}{\|x\|}\right) := \Sigma_0\left(\frac{x}{\|x\|}\right) w_0^*, \tag{4.8}$$

where $\Sigma_0(z) = I_n \otimes \sigma_0^\top(z) \in \mathbb{R}^{n\times nN_0}$, $z \in \mathbb{R}^n$ and $w_0^* = \mathrm{vec}\left(\left(W_0^*\right)^\top\right) \in \mathbb{R}^{nN_0}$. The norm matrix $\Sigma_0\left(\frac{x}{\|x\|}\right)$ has a finite upper-bound (in the matrix space) in view of (4.6), i.e.

$$\left\|\Sigma_0\left(\frac{x}{\|x\|}\right)\right\|_F \leq \sqrt{nN_0}, \ \forall x \in \mathbb{R}^n. \tag{4.9}$$

The identification problem can be understood as finding the vector of weights w_0 by implementing a weights adjustment law such that w_0 converges to w_0^* and x can be reproduced by $\hat{x}$,

where $\hat{x} \in \mathbb{R}^n$ represents the state vector of the DNN identifier. The following theorem introduces the first result for the convergence of the identification error.

Theorem 5. *Let assumptions 5-8 be satisfied and consider the approximation* (4.5) *for the homogeneous vector field f_0, with modeling error as in Remark 2 and the DNN identifier*[6] *given by*

$$\frac{d}{dt}\hat{x} = \|x\|^{\nu_0}\left[A\frac{\hat{x}}{\|x\|} + \Sigma_0\left(\frac{x}{\|x\|}\right)w_0 + \Omega K\Omega^\top e\right] + \sum_{i=1}^{m} f_i(x)\,u_i, \tag{4.10}$$

where $A \in \mathbb{R}^{n\times n}$ is a Hurwitz matrix, $e(t) = x(t) - \hat{x}(t)$ is the identification error (4.3)*, $w_0(t) \in \mathbb{R}^{nN_0}$ is the vector of weights adjusted as follows:*

$$\frac{d}{dt}w_0 = -\|x\|^{\nu_0} K\Omega^\top e, \tag{4.11}$$

and $\Omega \in \mathbb{R}^{n\times nN_0}$ is an auxiliary variable satisfying:

$$\tfrac{d\Omega}{dt} = \|x\|^{\nu_0}\left(A\frac{\Omega}{\|x\|} - \Sigma_0\left(\tfrac{x}{\|x\|}\right)\right). \tag{4.12}$$

If $K \in \mathbb{R}^{nN_0\times nN_0}$ is a positive definite symmetric matrix and the control inputs u in (4.1) *are such that $\exists x^- > 0$, $\exists x^+ > 0$, $x^- < \|x(t)\| < x^+ < +\infty$, $\forall t \in \mathbb{R}_+$ and the following persistent excitation (PE) condition*[7] *holds*

$$\int_t^{t+\ell_\omega} \Omega^\top(s)\Omega(s)ds \geq \vartheta_\omega I_{nN_0},\ \forall t \in \mathbb{R}_+, \tag{4.13}$$

for some $\ell_\omega > 0$ and $\vartheta_\omega > 0$. Then, there exist two class-$\mathcal{K}$ functions ρ_1 and ρ_2 such that:

$$\limsup_{t\to\infty} \|e(t)\| \leq \rho_1(\varepsilon_0), \tag{4.14}$$

$$\limsup_{t\to\infty} \|w_0(t) - w_0^*(t)\| \leq \rho_2(\varepsilon_0), \tag{4.15}$$

where ε_0 is given by (4.4).

[6] Notice that the DNN identifier (4.10) is not regular [30, 117] because it has a direct injection of the identification error e.

[7] Notice that in many cases the persistent excitation condition (4.13) can be fulfilled by a proper selection of a control input u (see e.g. [124] for more details about this).

Proof. To prove the ultimate boundedness of the identification error e and the input to state stability for the weights deviation variable $\tilde{\omega} := w_0^* - w_0$ consider the following auxiliary input:

$$\delta = e + \Omega\tilde{\omega}. \tag{4.16}$$

The dynamics of (4.16) satisfies:

$$\dot{\delta} = \dot{e} + \frac{d\Omega}{dt}\tilde{\omega} + \Omega\frac{d\tilde{\omega}}{dt}. \tag{4.17}$$

In (4.17), $\dot{e}$ corresponds to:

$$\dot{e} = \|x\|^{\nu_0}\left[\frac{Ae}{\|x\|} + \Sigma_0\left(\frac{x}{\|x\|}\right)\tilde{\omega} - \Omega K\Omega^\top e + d\left(\frac{x}{\|x\|}\right)\right]. \tag{4.18}$$

The time derivative of $\tilde{\omega}$ satisfies $\frac{d\tilde{\omega}}{dt} = -\frac{dw_0}{dt}$. Then, the substitution of (4.11), (4.12) and (4.18) on (4.17) yields:

$$\dot{\delta} = \|x\|^{\nu_0}\left[A\frac{e}{\|x\|} + \Sigma_0\left(\frac{x}{\|x\|}\right)\tilde{\omega} - \Omega K\Omega^\top e + d\left(\frac{x}{\|x\|}\right) + \left(A\frac{\Omega}{\|x\|} - \Sigma_0\left(\frac{x}{\|x\|}\right)\right)\tilde{\omega} + \|x\|^{\nu_0}\Omega K\Omega^\top e\right].$$

According to the definition of δ in (4.16), the previous equation is equivalent to:

$$\dot{\delta} = \|x\|^{\nu_0}\left(\frac{A}{\|x\|}\delta + d\left(\frac{x}{\|x\|}\right)\right). \tag{4.19}$$

Since A is Hurwitz, $\|x\| \neq 0$ and the modeling error $d(\cdot)$ corresponds to a bounded additive input (4.7) in (4.19), we can conclude that the auxiliary variable δ is also bounded, i.e.

$$\limsup_{t\to\infty} \|\delta(t)\| \leq \rho_0(\varepsilon_0), \tag{4.20}$$

where ρ_0 is a class-$\mathcal{K}$ function (see Lemma 1 in the Appendix). By the same result and considering (4.9), as well as (4.12), it is straightforward to observe that the variable Ω is also bounded, i.e.

$$\limsup_{t\to\infty} \|\Omega(t)\|_F \leq \bar{\Omega}, \tag{4.21}$$

where $\bar{\Omega}$ is depends on $\sqrt{nN_0}$ (see (4.9)), x^+ and x^-. Using the selected learning law (4.11) and (4.16), one obtains:

$$\frac{d\tilde{\omega}}{dt} = -\|x\|^{\nu_0}K\Omega^\top(\Omega\tilde{\omega} - \delta). \tag{4.22}$$

Due to boundedness of $\|x\|$ and Ω, the system (4.22) is input-to-state stable (ISS) with respect to the input δ (see Corollary 3 in the Appendix), i.e. the inequality (4.15) holds.

An estimate for the convergence quality of (4.3) can be obtained using (4.16). Since $e = \delta - \Omega\tilde{\omega}$, one may notice that:

$$\|e\| \leq \|\delta\| + \|\Omega\tilde{\omega}\|, \ \forall t \in R_+. \tag{4.23}$$

Using the norm relation in (4.23) (see e.g. [105]) and the fact that bound of $\tilde{\omega}$ depends on the upper bound for δ (see Lemma 1 in the Appendix), one gets:

$$\limsup_{t\to\infty} \|\Omega(t)\tilde{\omega}(t)\| \leq \limsup_{t\to\infty} \|\tilde{\omega}(t)\| \bar{\Omega} \leq \bar{\Omega}\rho_2(\varepsilon_0), \tag{4.24}$$

where ρ_2 is a class-$\mathcal{K}$ function from (4.15). The following relation for (4.23) takes place:

$$\limsup_{t\to\infty} \|e(t)\| \leq \rho_1(\varepsilon_0) := \rho_0(\varepsilon_0) + \bar{\Omega}\rho_2(\varepsilon_0). \tag{4.25}$$

□

4.3.2 The case of unknown control gains

In this section, we design a DNN identifier for the considered class of homogeneous system assuming that functions f_i are unknown, but admits the representation (4.5). Similarly to (4.8), we introduce the vectors w_i^* and the matrix-valued function Σ_i such that:

$$W_i^* \sigma_i \left(\frac{x}{\|x\|} \right) := \Sigma_i \left(\frac{x}{\|x\|} \right) w_i^*.$$

The matrices $\Sigma_i \left(\frac{x}{\|x\|} \right)$ have a finite upper bound (in the matrix space) in view of (4.6), i.e.

$$\left\| \Sigma_i \left(\frac{x}{\|x\|} \right) \right\|_F \leq \sqrt{nN_i}, \ \forall x \in \mathbb{R}^n. \tag{4.26}$$

Theorem 6. *Let assumptions 5-8 be satisfied and the control input u be selected as follows $u_i(t) = \frac{\tilde{u}_i(t)}{\|x(t)\|^{\nu_i - \nu_0}}, i = 1, \ldots, m,$, where $\tilde{u}_i : \mathbb{R}_+ \to \mathbb{R}$ are continuous uniformly bounded functions, i.e. that $|\tilde{u}_i(t)| \leq \tilde{U}, \forall t \in \mathbb{R}_+$ for some number $\tilde{U} > 0$. Consider the system* (4.1) *which can be*

represented in the form (4.5) *with an estimation error given in* (4.4)*. Define the DNN identifier as follows*

$$\frac{d\hat{x}}{dt} = \|x\|^{\nu_0-1}A\hat{x} + \|x\|^{\nu_0}\left[\sum_{i=0}^{m}\Sigma_i\left(\frac{x}{\|x\|}\right)w_i\tilde{u}_i + \Omega_i K_i \Omega_i^\top e\right],$$

where $\tilde{u}_0(t) := 1$ *,* $A \in \mathbb{R}^{n\times n}$ *is a Hurwitz matrix and* $e(t) = x(t) - \hat{x}(t)$ *is the identification error* (4.3)*,* $w_i(t) \in \mathbb{R}^{nN_i}$*,* $i = 0, 1, \ldots, m$ *are the vectors of the weights to be adjusted as follows*

$$\frac{d}{dt}w_i = -\|x\|^{\nu_0} K_i \Omega_i^\top e \tag{4.27}$$

and $\Omega_i \in \mathbb{R}^{n\times nN_i}$ *are auxiliary variables satisfying:*

$$\frac{d}{dt}\Omega_i = \|x\|^{\nu_0-1}A\Omega_i - \|x\|^{\nu_0}\tilde{u}_i\Sigma_i\left(\frac{x}{\|x\|}\right), \tag{4.28}$$

If $K_i \in \mathbb{R}^{nN_i\times nN_i}, i = 0, 1, \ldots, m$ *are positive definite matrices and the control inputs* u *in* (4.1) *are such that* $\exists x^- > 0$*,* $\exists x^+ > 0$*,* $x^- < \|x(t)\| < x^+ < +\infty$*,* $\forall t \in \mathbb{R}_+$ *and the following PE condition holds for all* $t \in \mathbb{R}_+$ *and some* $\vartheta_W > 0$ *and* $\ell_W > 0$*:*

$$\int_t^{t+\ell_W} G^\top(s)G(s)ds \geq \vartheta_W I_{nN_0}, \tag{4.29}$$

where the matrix $G \in \mathbb{R}^{n\times n\sum_{i=0}^{m}N_i}$ *is given by* $G = [\Omega_0, \Omega_1, \ldots, \Omega_m]$*. Then, there exist two class-*$\mathcal{K}$ *functions* ρ_1 *and* ρ_2 *such that*

$$\limsup_{t\to\infty}\|e(t)\| \leq \rho_1\left(\varepsilon^+\right), \tag{4.30}$$

$$\limsup_{t\to\infty}\|w_i(t) - w_i^*(t)\| \leq \rho_2\left(\varepsilon^+\right), \tag{4.31}$$

where $\varepsilon^+ = \max_{i=0,\ldots,m}\{\varepsilon_i\}$ *and* ε_i *are given by* (4.4)*.*

Proof. Consider $\tilde{\omega}_i = w_i^* - w_i$*,* $i = 0, 1, \ldots, m$ and the following auxiliary variable

$$\delta = e + \sum_{i=0}^{m}\Omega_i\tilde{\omega}_i. \tag{4.32}$$

Hence, the dynamics of (4.32) is

$$\dot{\delta} = \dot{e} + \sum_{i=0}^{m}\left[\frac{d\Omega_i}{dt}\tilde{\omega}_i + \Omega_i\frac{d\tilde{\omega}_i}{dt}\right], \tag{4.33}$$

where $\dot{e}$ is given by

$$\begin{aligned} \dot{e} &= \|x\|^{\nu_0-1}Ae \\ &+\|x\|^{\nu_0}\sum_{i=0}^{m}\left[\Sigma_i\left(\tfrac{x}{\|x\|}\right)\tilde{\omega}_i\tilde{u}_i-\Omega_i K_i\Omega_i^{\top}e\right]+p\left(x,u\right). \end{aligned} \tag{4.34}$$

The function $p\left(x,u\right)$ is defined in (4.4). The time derivative of $\tilde{\omega}_i$ satisfies $\frac{d\tilde{\omega}_i}{dt}=-\frac{dw_i}{dt}$. Then, the substitution of (4.27), (4.28) and (4.34) on (4.33) yields to:

$$\begin{aligned} \dot{\delta} &= \|x\|^{\nu_0-1}Ae+\|x\|^{\nu_0}\sum_{i=0}^{m}\left[\Sigma_i\left(\tfrac{x}{\|x\|}\right)\tilde{\omega}_i\tilde{u}_i-\Omega_i K_i\Omega_i^{\top}e\right]+p\left(x,u\right) \\ &+\sum_{i=0}^{m}\left(\|x\|^{\nu_0-1}A\Omega_i-\|x\|^{\nu_0}\tilde{u}_i\Sigma_i\left(\tfrac{x}{\|x\|}\right)\right)\tilde{\omega}_i+\|x\|^{\nu_0}\sum_{i=0}^{m}\Omega_i K_i\Omega_i^{\top}e. \end{aligned} \tag{4.35}$$

According to (4.32), (4.35) is equivalent to

$$\dot{\delta}=\|x\|^{\nu_0-1}A\delta+p\left(x,u\right). \tag{4.36}$$

Taking into account the identity $u_i(t)=\tilde{u}(t)/\|x\|^{\nu_i-\nu_0}$, from (4.4) we derive

$$\|p(x(t),u(t))\|\leq\|x\|^{\nu_0}\left(\varepsilon_0+\tilde{U}\sum_{i=1}^{m}\varepsilon_i\right).$$

Since A is Hurwitz and x is uniformly bounded from below and from above, we have (using Lemma 1)

$$\limsup_{t\to\infty}\|\delta(t)\|\leq\rho_0\left(\varepsilon^{+}\right), \tag{4.37}$$

for some class-$\mathcal{K}$ function ρ_0. Similarly, for (4.28) we conclude $\limsup_{t\to\infty}\|\Omega_i(t)\|\leq\bar{\Omega}_i$, where the number $\bar{\Omega}_i$ depends on $\sqrt{nN_i}$, $\tilde{U}$, x^{-} and x^{+} (see Lemma 1 in the Appendix). Using (4.27) and (4.32), we obtain:

$$\frac{d\tilde{\omega}_i(t)}{dt}=-\|x\|^{\nu_0}K_i\Omega_i^{\top}\left(\sum_{j=0}^{m}\Omega_j\tilde{\omega}_j-\delta\right). \tag{4.38}$$

or, equivalently $\frac{dW}{dt}=-\|x\|^{\nu_0}KG^{\top}(GW-\delta)$, where,

$$W(t)=\begin{bmatrix}\tilde{\omega}_0(t)\\ \tilde{\omega}_1(t)\\ \vdots\\ \tilde{\omega}_m(t)\end{bmatrix},\quad K=\begin{bmatrix}K_0 & 0 & \dots & 0\\ 0 & K_1 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & \dots & \dots & K_m\end{bmatrix}>0.$$

Due to boundedness of Ω_i, we derive (4.31)(From Corollary 3 in the appendix). Finally, using (4.32) we finish the proof:

$$\limsup_{t\to\infty} \|e(t)\| \leq \rho_1(\varepsilon^+) := \rho_0(\varepsilon^+) + \sum_{i=0}^{m} \rho_2(\varepsilon^+)\bar{\Omega}_i. \tag{4.39}$$

□

The complementary Corollary 3 and the associated Lemma 1 for the previous proofs are presented in the Appendix A.

The following Chapter presents the numerical results of the DNN homogeneous identifier, as well as the numerical results of the OC solution presented in the previous Chapters.

Chapter 5

Numerical Results

In this Chapter, the numerical simulations for the OC solution and the min-max DP for the class of uncertain system, are presented. The performance index using the proposed control laws is compared when another classical OCs are applied. The numerical results for both cases, the additive perturbation or uncertainty and the additive perturbation or uncertainty along with a uncertain multiplicative element associated with the input, are described. In addition, the effectiveness of the proposed identifier is verified by means of a simulation of a three-tank homogeneous model. In this example, the proposed identification scheme is compared with a non-homogeneous one.

The first section is devoted to the description of the numerical results for the OC solution, which was presented in the first Chapters of this thesis. The second section shows the numerical results of the DNN homogeneous identifier.

5.1 Numerical results of the min-max DP Approach

To evaluate the performance of the algorithm proposed in Chapter 2 and Chapter 3, an academic example for each case, describing a simplified nonlinear systems as in (2.1), was considered.

The simulations were made in Simulink Matlab® implementing a variable step integration method (minimum simulation step = 1 ms).

5.1.1 Numerical results for the case a)

The nominal matrices A and B for the system were:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -8 & -4 & -12 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 2.38 & 1.7 \end{bmatrix}.$$

The uncertain section of the model are is characterized by the matrices were:

$$F_a = \begin{bmatrix} sin(100*t) & sin(10000*t) & 0.6*sin(10000*t) \\ 0.8*cos(69*t) & 0.3*sin(854*t) & sin(90*t) \\ 0.5*sin(100t) & 0.4*cos(972t) & 0.6*cos(80t) \end{bmatrix},$$

$$G_a = \begin{bmatrix} 0.245*sin(10t) & cos(15t) \\ 0.157*cos(5t) & sin(0.3t) \\ sin(0.3t) & 0.245*cos(0.2t) \end{bmatrix}$$

The matrices F_a and G_a belong to the given set Ψ^{adm}. The boundary values for the uncertain matrices are: $\alpha_a = 1.00$ and $\beta_a = 0.13$. The proposed ANN was simulated using the following number of ANN weights $\omega = [\omega_{1b}\ \omega_{2b}\ \omega_{3b}\ \omega_{4b}\ \omega_{5b}\ \omega_{5b}]^\top$.

A simple validation scheme used the comparison of the states calculated if the robust sub-OC was evaluated in contrast to the numerical result achieved when the pole placement technique was proposed. In this numerical case, the desired vector of roots corresponded to

$$[-23.68,\ -0.15+0.8j,\ -0.15-0.8j]^\top,$$

that leads to the following control gain

$$K = \begin{bmatrix} 2.17 & -1.08 & -3.25 \\ -1.66 & -0.83 & -2.49 \end{bmatrix}.$$

The comparison of the state trajectories of both controllers (the ANN based sub-optimal and the pole placement based) appears in Figure 5.1.

This comparison demonstrates that all the states converge to a bounded region near the origin with an upper bound of 0.03 after 1.0 seconds. However, the trajectories of the system controlled with the sub-optimal controller approaches the origin slower than the ones controlled with the pole-placement method.

The trajectories depicted in Figure 5.1 served to estimate the norm of x. The Euclidean norm was used for comparison purposes. This comparison confirmed that the pole-placement technique enforces a faster movement of the state trajectories toward the origin within the first 1.5 seconds. The variation of the state norm also served to estimate the variation of the VF, as well as the functional time evolution (Figure 5.2).

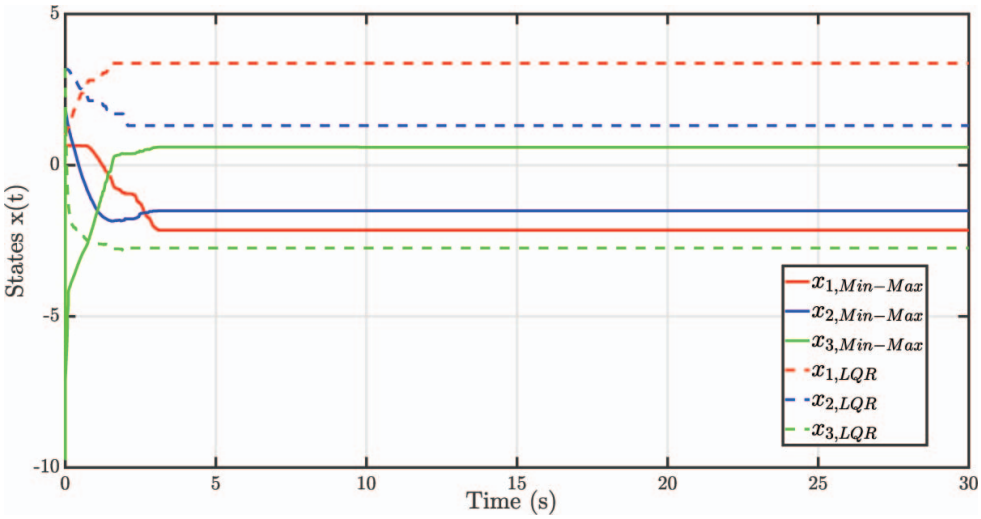

Figure 5.1: States determined with the states and control actions using the sub-optimal and pole-placement approaches.

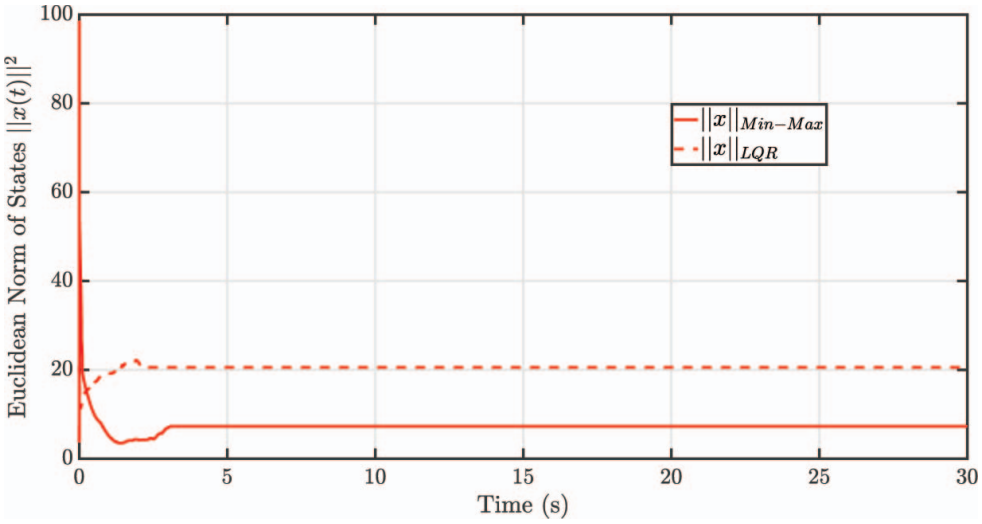

Figure 5.2: Norm of the states vector estimated with the states and control actions using the sub-optimal and pole-placement approaches.

The evolution of the states and the evaluation of the control function yield to determine the variation of the performance index. This comparison was a major element to define the advantage of the control design attained in this thesis. The comparison of the performance index calculated with the sub-OC and the pole-placement technique showed a significant reduction of such a value after 0.05 seconds and a final (after 0.5 seconds) of 33 $\%$ after 0.5 seconds (Figure 5.3).

The depicted results also show that the variation of the initial condition for the weights in the ANN provides a significant variation in the performance index temporal evolution. Also, with the local optimal values for the weights of the ANN, the lowest performance index is attained (Figure 5.3).

With the aim of showing the dependence of the performance index with respect to the initial condition of the weights, Figure 5.4 demonstrates the value of the performance index $J(u(\cdot))$ calculated with $T = 30s$. This behavior confirms the usefulness of the suggested algorithm, which is in charge of adjusting the ANN weights.

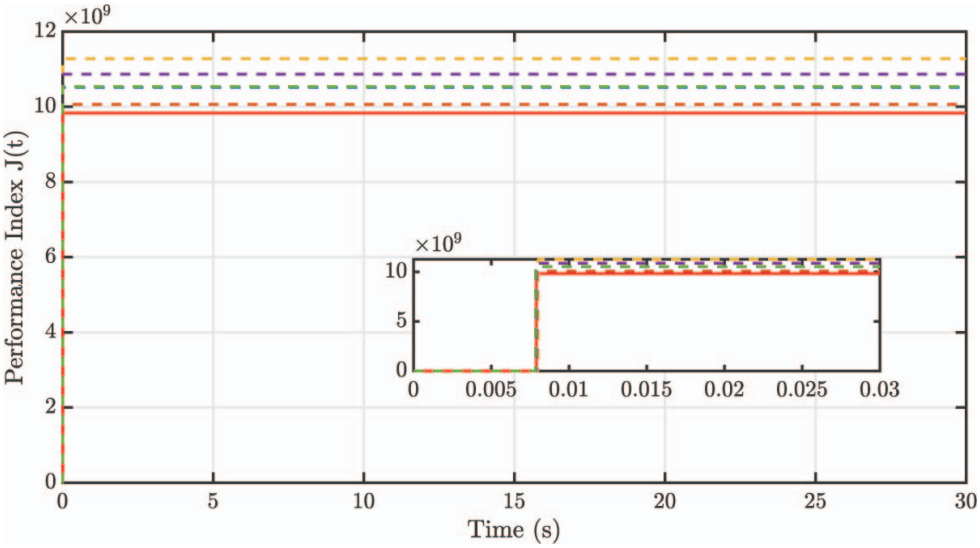

Figure 5.3: Cost functional estimated with the states and control actions using the sub-optimal and pole-placement approaches.

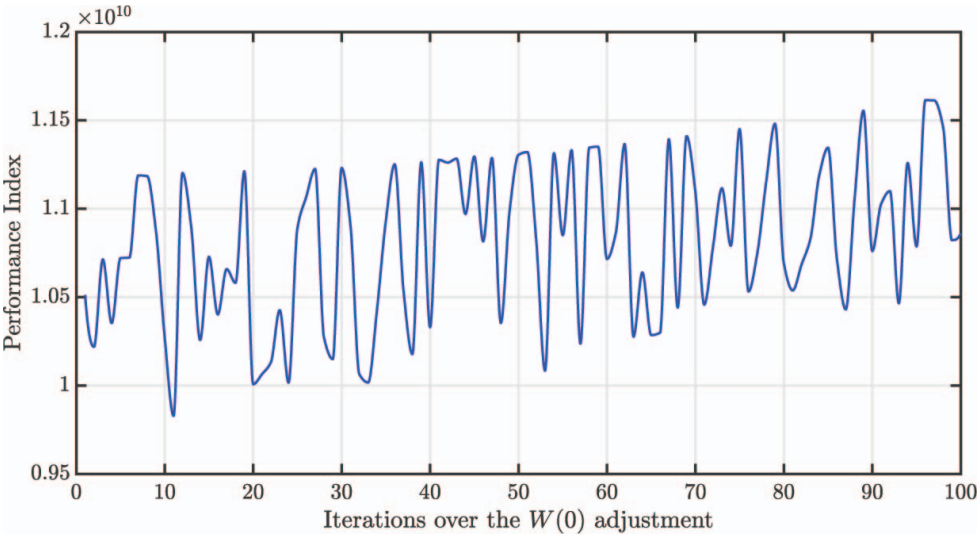

Figure 5.4: Cost functional estimated with the states and control actions using the sub-optimal approach presented as a function of the variation of the initial weights in the ANN.

Figure 5.5 shows the variation of the approximate *Robust* VF V_a obtained with the states and the control signal calculated from the functional $J(u(\cdot))$. The decreasing behavior of this VF confirmed the applicability of the approximation based on the ANN, as well as the numerical

methodology used to adjust the initial weights (Chapter 3, Section 2).

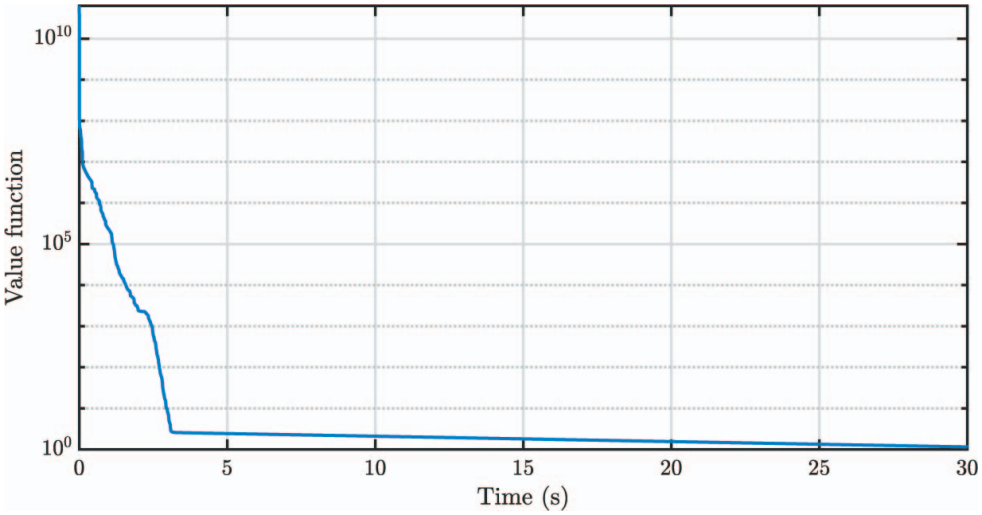

Figure 5.5: Value function calculated with the states and control actions using the sub-optimal controller obtained by the approximation based on neural networks.

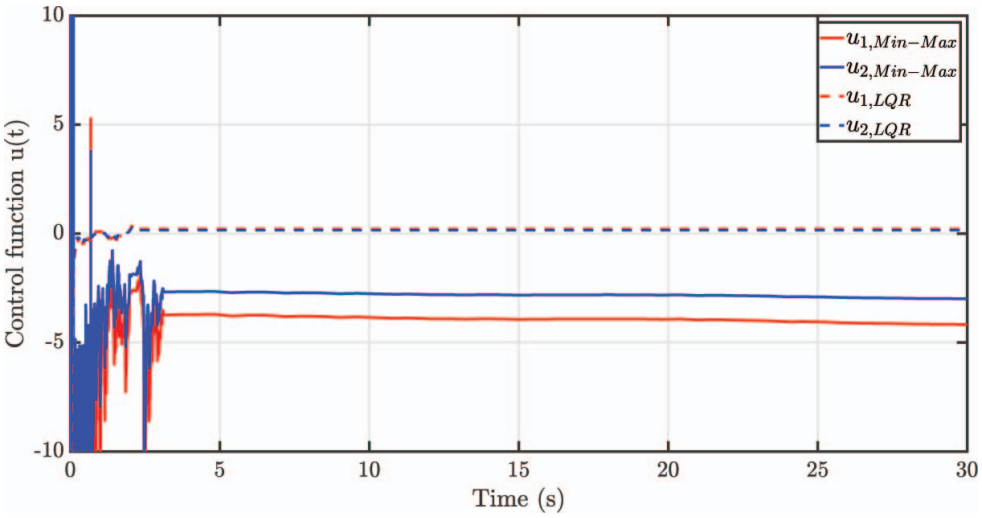

Figure 5.6: Comparison of the control signals obtained by the sub-optimal approach and the PD design strategy.

Figure 5.6 demonstrates the time evolution of the controller using the sub-optimal solution and the pole-placement approach. The sub-optimal control signals were smaller than those produced by the pole-placement except during the first 0.05 seconds. This fact justifies the

differences between the convergences of the states to a bounded zone near the origin. It is relevant that sub-optimal control signals reach the values obtained when the pole-placement controller is considered. This fact can be confirmed after the first second of simulation.

Figure 5.7 depicts the variation for the norm associated to the controller signal u. The time evolution of this function was considered due to its participation on the VF, as well as the functional estimation. The variation of the control norm together with the variation of the states norm justifies the increment of the function when the sub-OC is considered in comparison to the regular pole-placement method, which is not taking care of the perturbations effect.

Figure 5.8 shows the variation of all the four components in the weights vector $\tilde{\omega}$ that were considered in the approximation of the VF solution for the HJB equation $V_a(t,x)$. Notice that in opposition to the usual behavior in the ANN approximation of uncertain functions, the weights obtained in this study do not converge to constant values because of the influence of the system states x and the time dependent solution of the Riccati equation given by P.

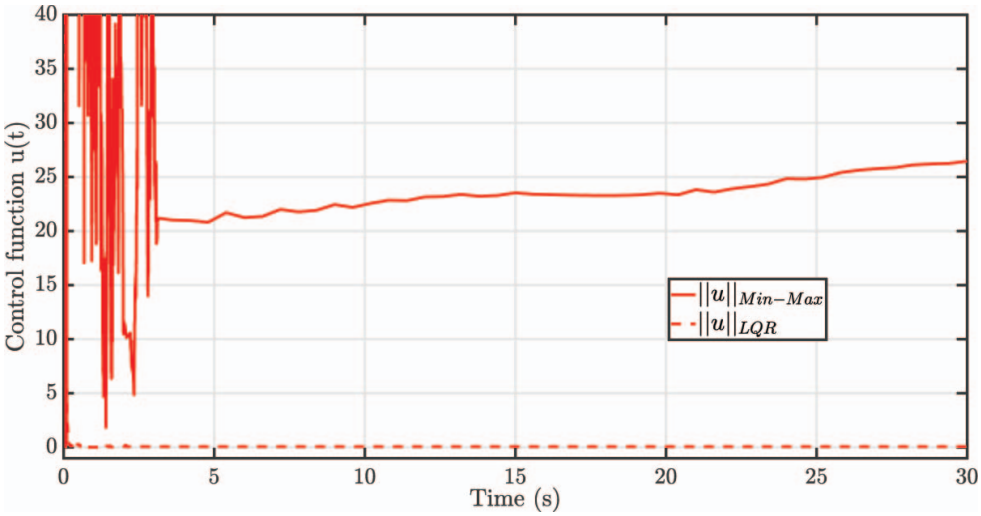

Figure 5.7: Norm of control signal

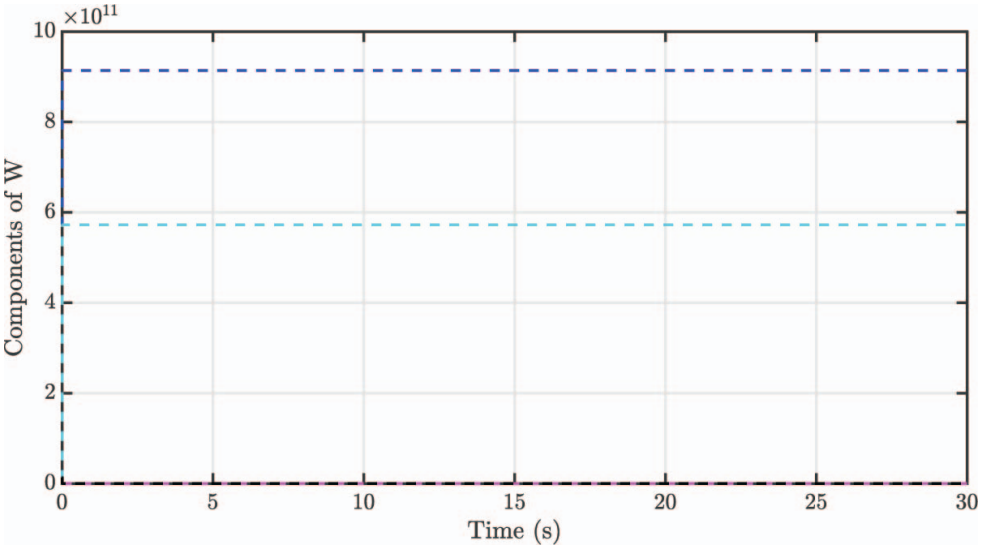

Figure 5.8: Time dependent trajectories of $\omega_b(t)$.

The evolution of the weights included in the ANN structure can be complemented with the variation of sigmoidal functions. Their variations demonstrate the effect of the states evolution over the ANN structure (Figure 5.9).

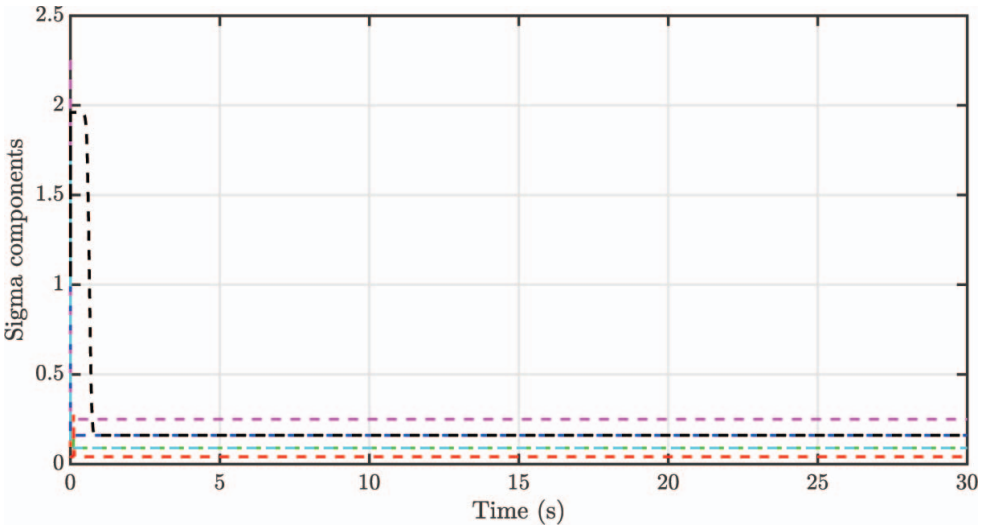

Figure 5.9: Time dependent trajectories of $\sigma_b(t)$.

Figure 5.10 shows the time evolution of all the four elements included in P which was obtained by the on-line numerical solution for the Riccati time varying matrix differential equation.

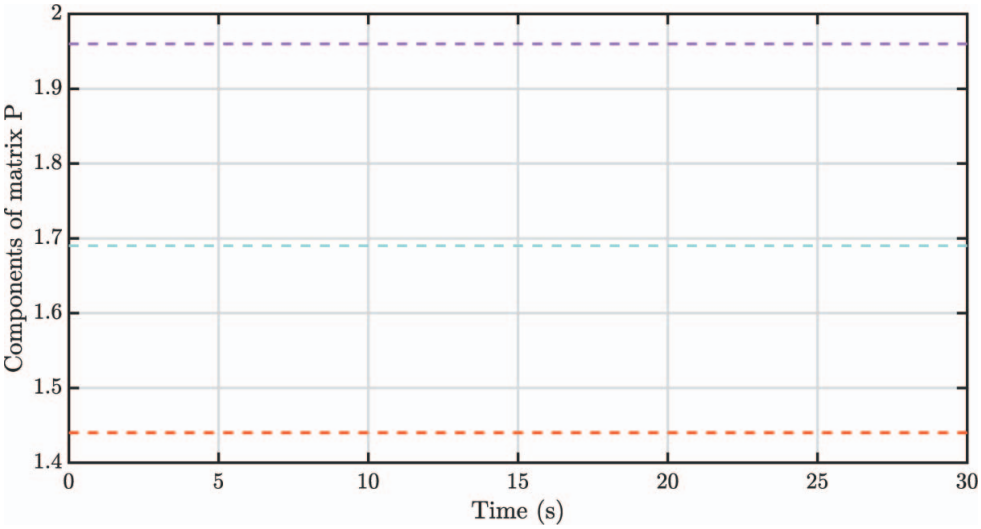

Figure 5.10: Time evolution of the components of the matrix P calculated by the numerical solution of the time-dependent Riccati matrix differential equation.

The ANN approximation proposed in this thesis represents a contribution to the robust realization of OC for systems with the admissible class of parametric uncertainties and external perturbations.

The proposed numerical results demonstrated the existence of the practical stable equilibrium point associated to the origin. This result justifies the theoretical result associated to the stability analysis based on the Lyapunov-like function (energetic type) (Chapter 3, Section 4). Moreover, the approximate solution based on the Riccati matrix equation also shows the convergence to bounded and constant values. Interestingly, the weights of the ANN are not converging to constant values in contrast to the regular ANN approximations of the HJB for the optimal control realization.

The existing approximate results of the OC for the class of systems considered in this study have not proposed the analytic results attained here. Notice that this study provides the evaluation of the approximate function over the sub-OC, as well as its impact on the Hamiltonian associated to the ANN approximation. The quasi-linear form of the uncertain system has motivated the proposal of a mixed controller using a linear form plus the approximated solution based on the ANN.

5.1.2 Numerical results for the case b)

The system as in (2.1), is characterized with the following matrices

$$A = \begin{bmatrix} 0 & 1 \\ 0.2 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}, \quad F_b = \begin{bmatrix} 100\sin(100x_1) & 5\sin(10000x_1) \\ -20\cos(10000x_2) & -6\cos(69x_2) \end{bmatrix},$$

$$G_b = \begin{bmatrix} 0.245\sin(0.1x_1) & 0.24\sin(50x_1) \\ -0.25\cos(15x_1) & -0.2\cos(x_2) \end{bmatrix},$$

The initial condition for (2.1) is $x_0 = \begin{bmatrix} 5, & 15.5 \end{bmatrix}^\top$. The cost functional that must be optimized satisfies (2.3). The matrices Q and R for the cost functional are

$$Q = \begin{bmatrix} 120 & 0 \\ 0 & 130 \end{bmatrix}, \quad R = \begin{bmatrix} 0.04 & 0 \\ 0 & 0.028 \end{bmatrix}.$$

The final time is selected as $T = 20$ seconds. The ANN structure has four activation functions with the following parameters:

$$y_1 = \begin{bmatrix} 0.001 \\ 0.004 \end{bmatrix}, \; y_2 = \begin{bmatrix} 0.007 \\ 0.003 \end{bmatrix}, \; y_3 = \begin{bmatrix} 0.003 \\ 0.002 \end{bmatrix}, \; y_4 = \begin{bmatrix} 0.004 \\ 0.001 \end{bmatrix}, \tag{5.1}$$

$$\theta_1 = 0.02, \; \theta_2 = 0.03, \; \theta_3 = 0.04, \; \theta_4 = 0.05.$$

The initial condition to evaluate the dynamics of the matrix $P(t)$ is selected as

$$P(0) = \begin{bmatrix} 43.75 & 3.75 \\ 3.75 & 28.75 \end{bmatrix}. \tag{5.2}$$

The initial weights are obtained with Algorithm 1 after 222 iterations, such that,

$$w_{222} = \begin{bmatrix} 0.8, & 1.6, & 3.2, & 4 \end{bmatrix}^\top$$

.

The designed controller was compared with a Linear Quadratic Regulator (LQR), whose gain is computed using the nominal matrices A and B. The estimated gain matrix (using the Matlab®

function LQR) is

$$K_{LQR} = \begin{bmatrix} 12.0493 & -51.5647 \\ -66.5233 & -15.4592 \end{bmatrix}.$$

In Figure 5.11, the comparison between the cost functional obtained with the ANN controller and the cost functional obtained with the LQR is depicted. The smaller functional is result of the system in closed loop using the ANN approximation for the VF.

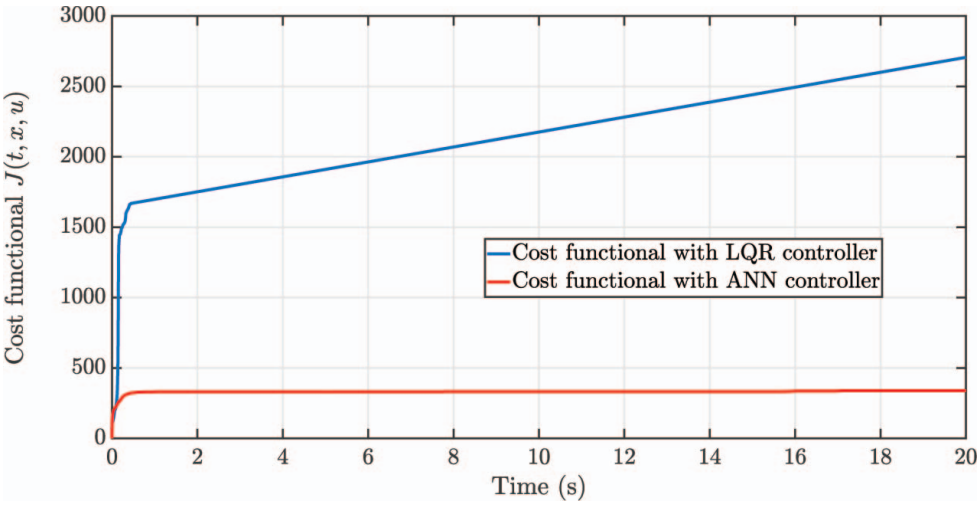

Figure 5.11: Comparison between the cost functionals obteined with the designed controller and with the LQR.

Figure 5.12 shows the evolution of the system states using both controllers, the LQR and the ANN sub-OC. The first subplot of Figure 5.12 shows that the first state $x_1(t)$ reaches the origin in less than two seconds. However, the first state $x_1(t)$ obtained with the LQR implementation converges only to a zone around the origin with size of 0.44. In the second subplot, similar results for the second state x_2 are presented with the zone having the size of 0.14.

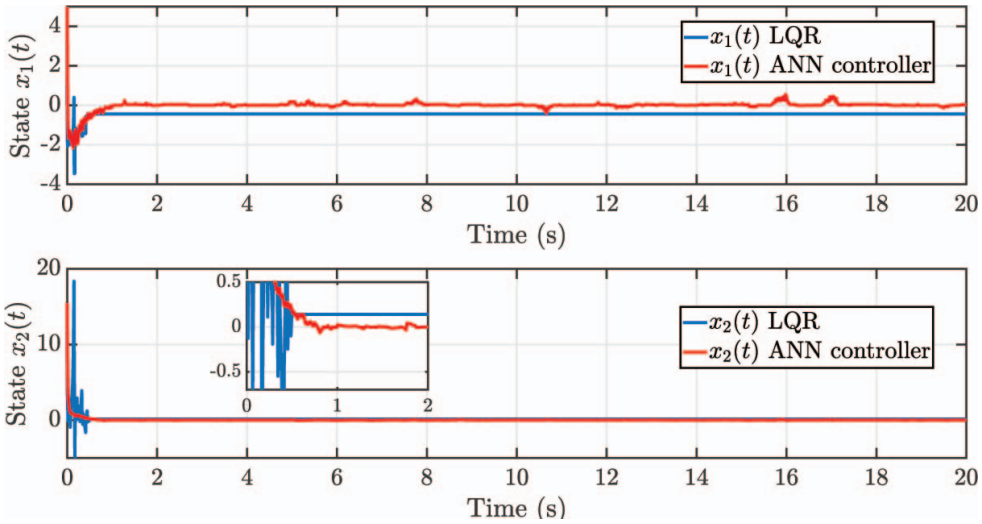

Figure 5.12: Evolution of the states with the ANN controller and with the LQR.

Figure 5.13 depicts the evolution of the elements of matrix $P(t)$. The elements converge to constant values around two seconds. The matrix remains symmetric and positive definite during the simulation time.

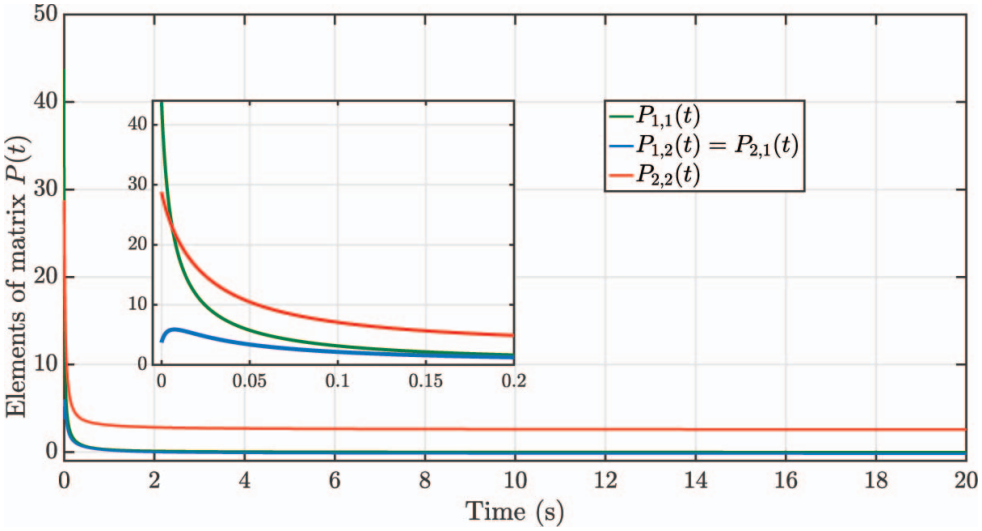

Figure 5.13: Evolution of the elements of the matrix $P(t)$ (Solution of the Riccati Equation).

The evolution of the ANN weights (Figure 5.14) shows that their values converge to constant values in less than one second.

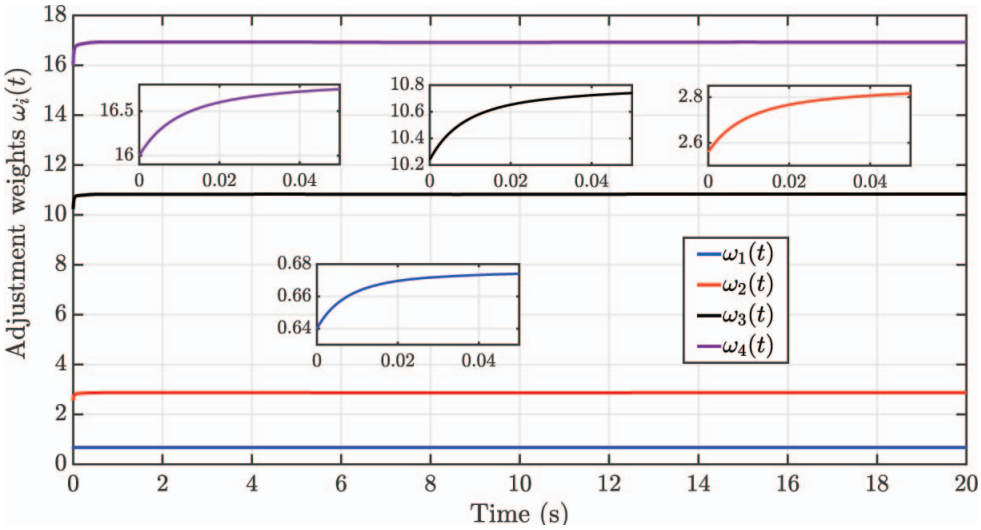

Figure 5.14: Evolution of the adjustment weights $\omega_i(t)$.

In Figure 5.15, the control signals with the LQR and the ANN controller are showed. The obtained signals with the LQR present oscillations during the first 0.5 seconds, while the ANN based controller does not enforce similar oscillations.

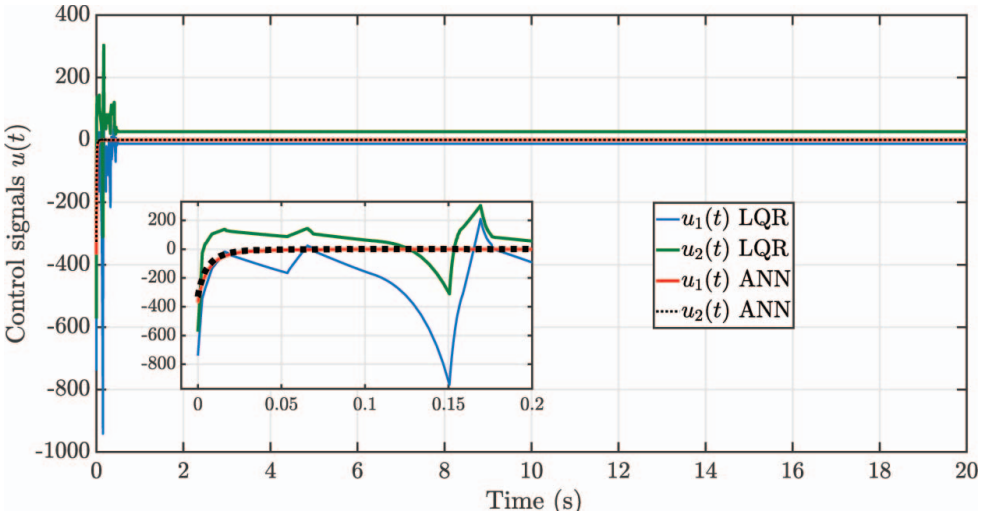

Figure 5.15: Evolution of the control signals $u(t)$.

The time varying changes of the four activation functions evaluations are depicted in Figure 5.16. All the activation functions are selected as the square of (1.15) with the parameters proposed in (5.1).

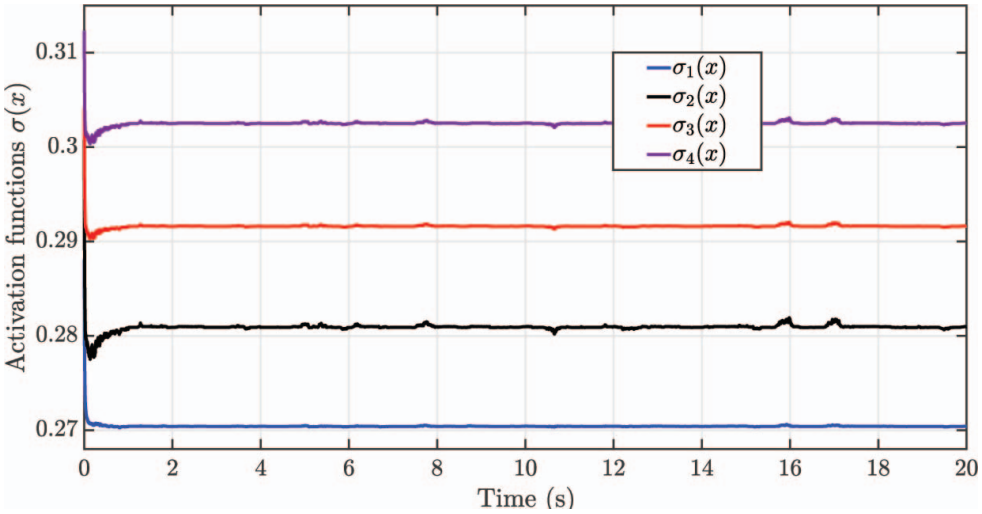

Figure 5.16: Evolution of the activation functions

In Figure 5.17, the Euclidean norm of the states with both controllers, the LQR and the ANN controller, is compared. This result shows that the *robust* optimal solution has smaller values of this norm after 20 seconds. Indeed, the states converge to the origin around two seconds yielding the norm also to the origin at the same time.

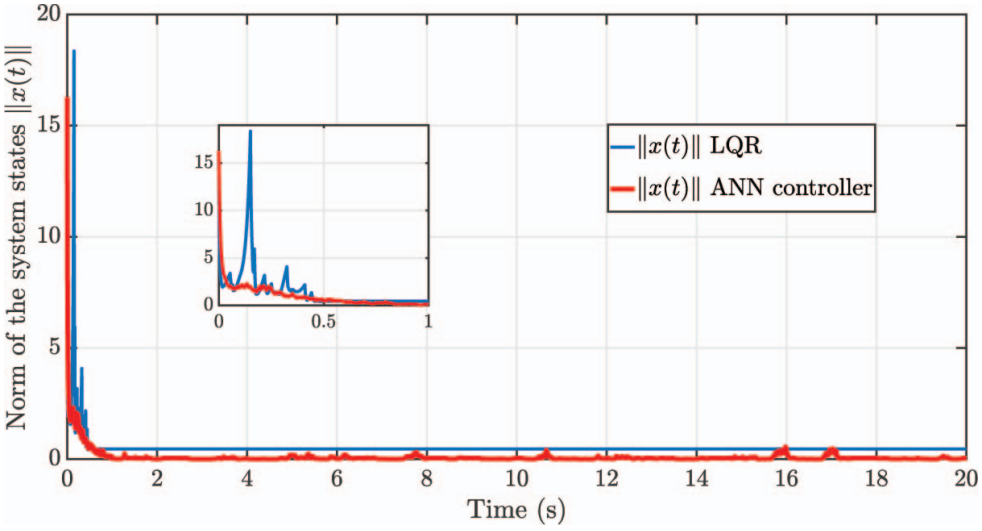

Figure 5.17: Norm of the states of the systems

Figure 5.18 depicts the comparison between the Euclidean norm of the control signal with the LQR and the Euclidean norm with the suboptimal controller. The time evolution of the Euclidean

norm of the controller with the ANN is smaller than the obtained with the LQR. This result confirms that smaller states norm provided by the ANN based controller does not imply a higher control realization.

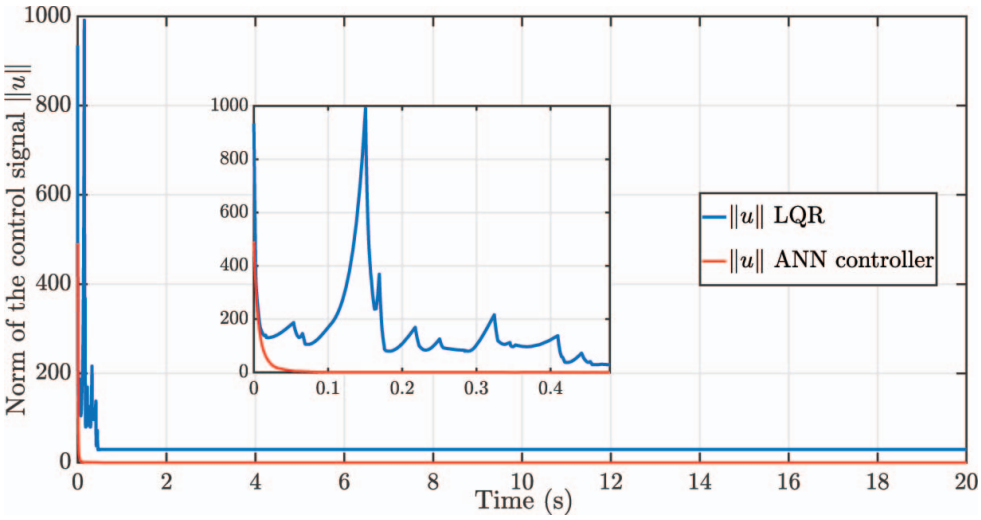

Figure 5.18: Norm of the control input

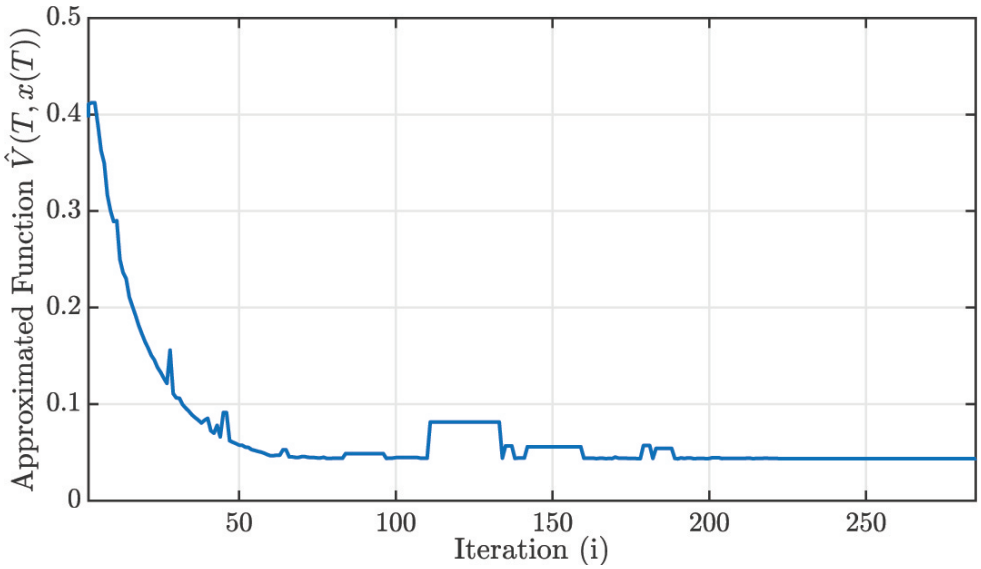

Figure 5.19: Final value of the approximated function $\hat{V}(T, x(T))$

Figure 5.19 shows the final value evolution of the approximated ANN solution (approximated VF). The measured changes of the final value are consequence of the adjustment of the initial

condition for the weights of the ANN using Algorithm 1. This result confirms the convergence of the proposed algorithm to adjust the initial weights of the ANN approximation. Notice that despite the presence of local minimum values, the algorithm adjusts the initial weights values until getting an invariant value (0.047) observed after 200 iterations.

The following section describes the numerical results for the other application of the ANN approximation capabilities, the DNN homogeneous identifier.

5.2 DNN homogeneous identification for a three tank system

To show the performance of the designed identifier, we use a three-tank system. The mathematical model for the system is presented with its homogeneity degrees, however the explicit model is not used in the application of the identifier.

Consider the three-tank system (Figure 5.20).

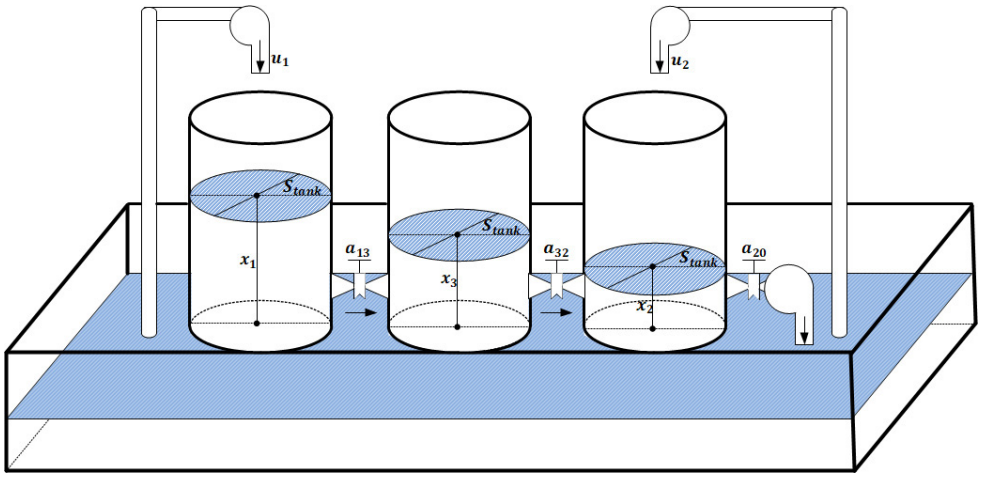

Figure 5.20: Three tank system.

The three-tank nonlinear system admits the following dynamic model [97, 125, 126]:

$$\begin{aligned}
\dot{x}_1 &= S_{tank}^{-1}\left(-a_{13}\lceil x_1 - x_3\rfloor^{0.5} + u_1\right), \\
\dot{x}_2 &= S_{tank}^{-1}\left(a_{32}\lceil x_3 - x_2\rfloor^{0.5} - a_{20}\lceil x_2\rfloor^{0.5} + u_2\right), \\
\dot{x}_3 &= S_{tank}^{-1}\left(a_{13}\lceil x_1 - x_3\rfloor^{0.5} - a_{32}\lceil x_3 - x_2\rfloor^{0.5}\right),
\end{aligned} \tag{5.3}$$

where $x_1,[L]$, $x_2,[L]$ and $x_3,[L]$ represent the liquid level of each tank respectively, $S_{tank},[m]$ is the diameter of the three tanks, the input flows $u_1,[L\cdot s^{-1}]$ and $u_2,[L\cdot s^{-1}]$ are the control signals and the constant parameters a_{13}, a_{32} and a_{20} are coefficients related with the outflow rate according to Torricelli's rule.

Table 5.1: Parameters for the simulation of the three-tank system

Parameter	**Description**	**Value**
S_{tank}	Tank diameter	$1m$
a_{13}	Outflow rate coefficient	$3\ m^{\frac{3}{2}}/s^2$
a_{32}	Outflow rate coefficient	$2\ m^{\frac{3}{2}}/s^2$
a_{20}	Outflow rate coefficient	$1\ m^{\frac{3}{2}}/s^2$
x_0	Initial conditions	$[3,\ 1,\ 2]^\top$

Obviously, the system admits the representation (4.1) with the functions vectors f_i

$$f_1(x)=\begin{bmatrix}1\\0\\0\end{bmatrix},\ f_2(x)=\begin{bmatrix}0\\1\\0\end{bmatrix},\ f_0(x)=\frac{1}{S_{tank}}\begin{bmatrix}-\theta_1\lceil x_1-x_3\rfloor^{0.5}\\ \theta_3\lceil x_3-x_2\rfloor^{0.5}-\theta_2\lceil x_2\rfloor^{0.5}\\ \theta_1\lceil x_1-x_3\rfloor^{0.5}-\theta_3\lceil x_3-x_2\rfloor^{0.5}\end{bmatrix},\ x=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}.$$

which are homogeneous with the associated degrees $\nu_0=-0.5$, $\nu_1=\nu_2=0$. We assume that the functions f_i are unknown, but their homogeneity degrees are known.

The matrix A for the identification algorithm has been selected as follows:

$$A=\begin{bmatrix}-5.1 & -3.5 & -3\\ -22.5 & -32 & -17\\ -4 & -2 & -12\end{bmatrix},$$

and the activation functions are given by (1.15) with $N_i = 3$. The constant parameters $\theta_{0,1} = 5$, $\theta_{0,2} = 10$ and $\theta_{0,3} = 20$, the constant vectors are:

$$y_{0,1} = \begin{bmatrix} 0.01 \\ 0.02 \\ 0.03 \end{bmatrix}, \; y_{0,2} = \begin{bmatrix} 0.01 \\ 0.04 \\ 0.01 \end{bmatrix}, \; y_{0,3} = \begin{bmatrix} 0.04 \\ 0.01 \\ 0.06 \end{bmatrix}.$$

Equal constant parameters $\theta_{i,j}$ and $\theta_{0,j}$ are chosen $(\theta_{i,j} = \theta_{0,j})$, as well as in the case of the vectors $y_{i,j} = y_{0,j}$. The initial conditions for the adjustment laws are selected as:

$$\begin{gathered} w_0(0) = \begin{bmatrix} 1 & 2 & 1 & 1 & 3 & 2 & 1 & 2 & 3 \end{bmatrix}^\top \quad w_1(0) = \begin{bmatrix} 1 & 3 & 1 & 5 & 3 & 2 & 1 & 2 & 3 \end{bmatrix}^\top \\ w_2(0) = \begin{bmatrix} 1 & 4 & 1 & 2 & 1 & 6 & 1 & 3 & 3 \end{bmatrix}^\top, \end{gathered} \tag{5.4}$$

and the gain matrices $K_0 = I_3 \otimes \tilde{K}_0$ and $K_1 = K_2 = I_3 \otimes \tilde{K}_1$, with:

$$\tilde{K}_0 = \begin{bmatrix} 5220 & -1044 & 1566 \\ -1044 & 1566 & 522 \\ 1566 & 522 & 7830 \end{bmatrix}, \tilde{K}_1 = \begin{bmatrix} 5250 & -1050 & 1575 \\ -1050 & 1575 & 525 \\ 1575 & 525 & 7875 \end{bmatrix}.$$

The performance of the designed algorithm was compared with a classical DNN identifier (see [30], Chapter 2). The numerical simulations for the identifiers are made in Simulink Matlab® by using the Runge Kutta integration method with a step of 0.1 ms.

For the classical DNN identifier, the following parameters are considered, the matrix A is the same as the selected for the homogeneous one, the constants for the vector of activation functions associated to f_0 in the classical DNN identifier were the same as the parameters for $\sigma_0(x)$ in the homogeneous identifier. The matrix of activation functions $\phi : \mathbb{R}^3 \to \mathbb{R}^{3\times 2}$ is selected as $\phi(x) = [\sigma_1(x), \sigma_2(x)]$. The initial conditions for the adjustment law are:

$$W_1(0) = \begin{bmatrix} 10 & 23 & 12 \\ 14 & 33 & 26 \\ 14 & 25 & 36 \end{bmatrix}, \; W_2(0) = \begin{bmatrix} 13 & 34 & 15 \\ 56 & 53 & 2 \\ 12 & 23 & 35 \end{bmatrix}$$

The initial conditions are the same for both identifier algorithms $\hat{x}(0) = [5, 2, 3]^\top$.

In Figures 5.21, 5.22 and 5.23, the signal with the displayed name *DNN classical Identifier* (dotted line, red color) describes the states for the identifier using the algorithm presented in [30]. The states of the designed identifier for homogeneous systems that assumes all the functions f_i as unknown, appear represented with the reference *DNN Identifier for homogeneous systems* (dashed line, blue color).

In Figure 5.21, the first estimated state of the identifier for homogeneous systems converges faster (before 0.01 seconds) to the state of the uncertain system (solid line, black color) than the classical DNN identifier. The estimation of the classical algorithm contains oscillations of relevant amplitude with respect to the values of the states.

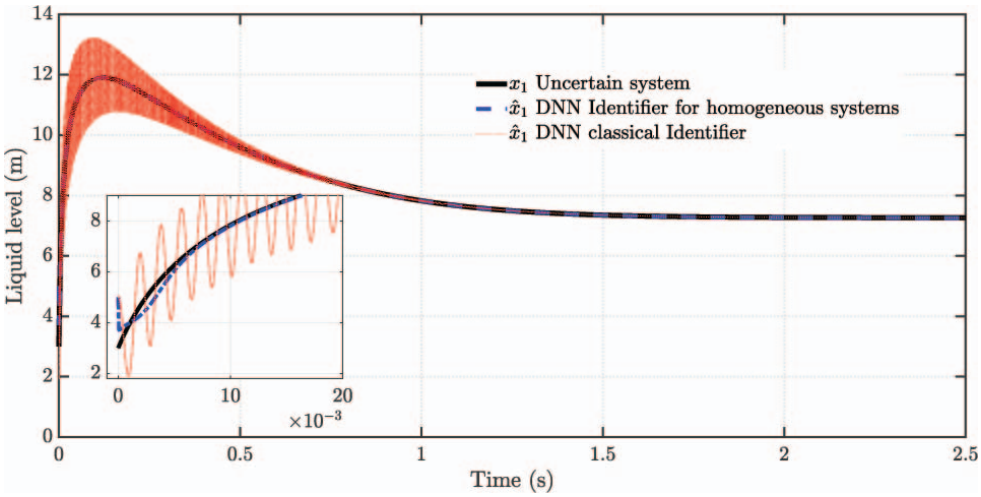

Figure 5.21: Identification result for the first state (x_1).

Figure 5.22 shows similar results for both identifiers. In the closer view, it is possible to notice that the second estimated state (dotted line, color red) of the identifier using a classic series-parallel DNN structure presents bigger oscillations and has a slower convergence than the homogeneous algorithm. The second estimate state (dashed line, blue color) of the identifier for homogeneous systems converges in an approximated time of 8 ms.

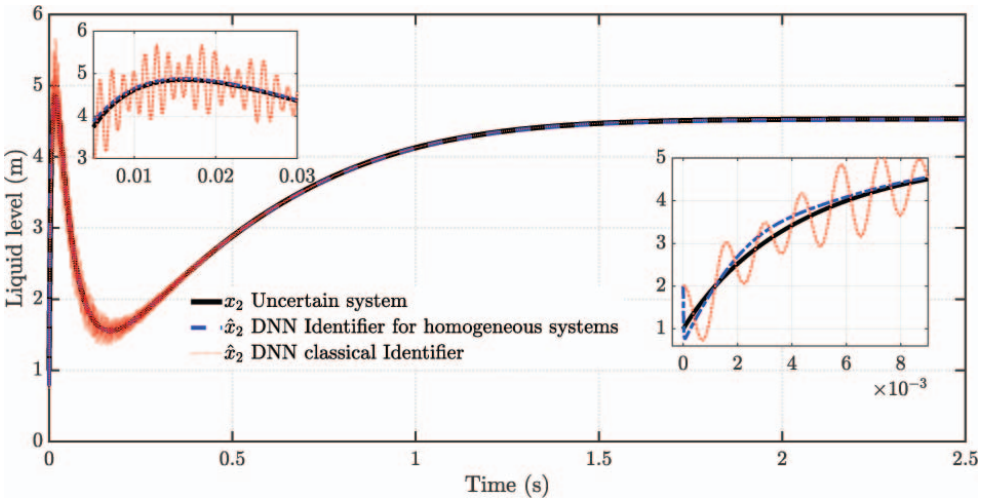

Figure 5.22: Identification result for the second state (x_2).

In Figure 5.23, the estimation for the third state of the three-tank plant is depicted. The estimate state convergence using the algorithm devoted to homogeneous systems (before 0.02 seconds) is faster than the convergence of the estimate state with the classical DNN identifier (around 0.3 seconds). In addition, the estimation with the classical algorithm presents oscillations.

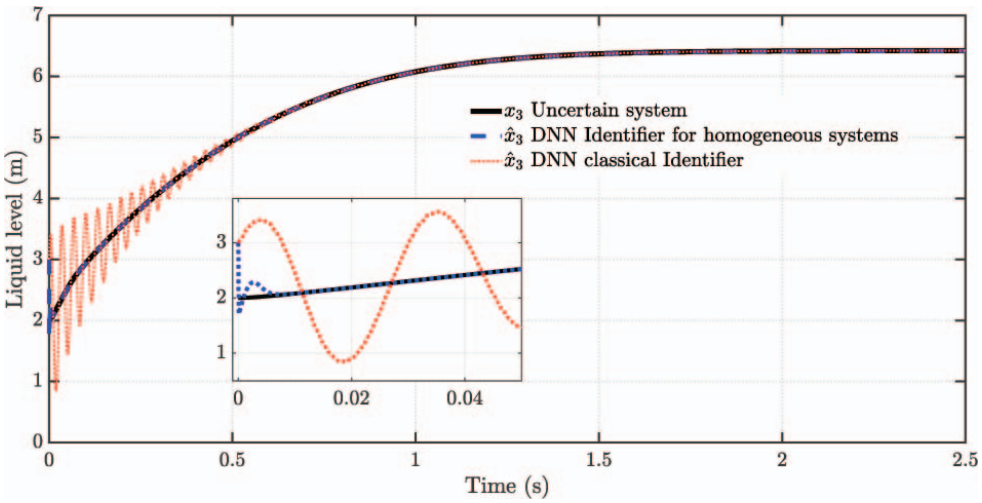

Figure 5.23: Identification result for the third state (x_3).

In Figure 5.24, the comparison for the norm of the identification error with the classic DNN structure identifier (dotted line, color red) and the DNN homogeneous identifier for systems with

uncertain model (solid line, blue color) is depicted. The identification error with the classical DNN identifier converges around 0.8 seconds versus 0.02 seconds of the identification error when the DNN identifier is devoted to homogeneous systems.

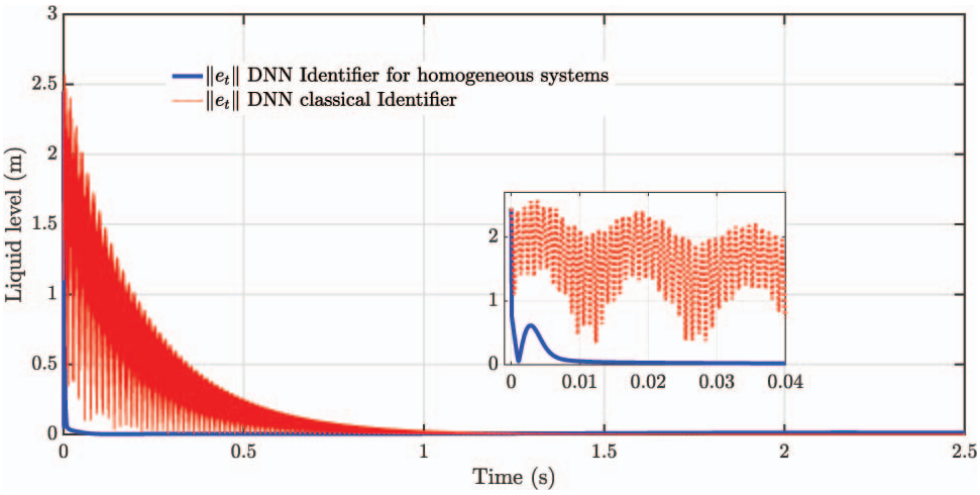

Figure 5.24: Norm of the identification error.

For the simulation, considering all the vector functions unknown used the same 9 activation functions for both algorithms. In the case of the classical DNN structure, the activation functions matrix for the terms associated to the input has the same elements of the two activation functions vectors of the algorithm for homogeneous systems.

General conclusions

In this thesis, the finite time horizon OCP for a class of uncertain system has been tackled by proposing a class of a min-max sub-optimal control with a NDP approach. The controller used an ANN approximation of the VF. The proposed approximated solution of the HJB equation has been based on the given ANN structure added with a classical quadratic form of the state weighted by a time-dependent positive definite matrix. The ANN structure obeys the nature of the VF by using quadratic terms and by choosing sigmoidal activation functions. The tuning of the free-parameters for the approximation adjustment was solved with a recurrent numerical algorithm based on the application of a Levenberg-Marquardt algorithm.

Compared with other works, in this thesis, the effect of the unknown terms is added in the main theorems, and the structure of the ANN considered the natural quadratic structure of the nominal part of the system. The numerical simulation of the proposed controller was implemented to regulate a quasi-linear system with parametric uncertainties and external perturbations and it was compared with a classical pole-placement controller. The performance of the sub-optimal controller was illustrated by the numerical simulation proposed in this work.

The ANN approximation proposed in this thesis represented a contribution to the robust realization of OCs for systems with the admissible class of parametric uncertainties and external perturbations considered in this study.

The proposed numerical results demonstrated the existence of the practical stable equilibrium point associated to the origin. This result justified the theoretical result associated to the stability analysis based on the Lyapunov-like function (energetic type). Moreover, the approximated solution based on the Riccati matrix equation also showed the convergence to bounded

and constant values.

The existing approximate results of the OC for the class of systems considered in this study have not proposed the analytic results attained here. Notice that, this study provided the evaluation of the approximate function over the sub-optimal control, as well as its impact on the Hamiltonian associated to the ANN approximation. The quasi-linear form of the uncertain system has motivated the proposal of a mixed controller using a linear form plus the approximated solution based on the ANN.

The results were obtained for two cases of the considered class of system. The first case considered an additive unknown term and the second case considered an additional uncertain multiplicative element associated with the input. In both structures the unknown elements can represent model uncertainties or bounded perturbations.

On the other hand, the approximation capabilities of the ANN have been used for the design of an identifier for a class of homogeneous systems.

The design of an identification algorithm for standard homogeneous uncertain control systems was developed in this thesis using a DNN approach and stability Lyapunov theory to derive the learning laws.

The convergence of the identification error and the learning laws is proven theoretically, given as a result the ultimate boundedness and ISS for the identification error and the adaptive weights of the network.

The PE condition for the designed identifier is used in the stability analysis, in other words, a constructive form of the PE condition applied for the realization of the convergence analysis is presented.

Numerical simulations were presented using a three-thank homogeneous model to demonstrate the performance of the identification algorithm. The homogeneous identifier was compared with a classical DNN identifier. The results of such comparison demonstrated that the homogeneous identifier approximates the homogeneous model of the three-tank simulation with less oscillations and a smaller identification error.

Future work

After the realization of this thesis, it is planned to extend the applications of the solutions obtained here, for example:

- To implement the OC solution in an experimental platform and to study the performance index evolution using different numbers of neurons in the ANN structure.
- In addition, the design of other class of algorithms for the learning laws to achieve an approximation in a predefined exponential time or learnings laws with finite time convergence. This aims are for its used in the identification problem.
- As a future extension dor the identifier, it is planned the study of the particular different cases based on the sign of the homogeneous degree and its identification due to the requirements of Assumption 6, as well as the implementation in an experimental platform.
- Other future objective is to extend the result of the identifier for generalized homogeneity and to use it for the design of the OC solution for the problem presented in this thesis but considering an generalized homogeneous structure in the system.

Bibliography

[1] R. M. Rilke and F. Bermúdez C., *El libro de horas: Edición Bilingüe*. Hiperión, 2005. 5

[2] M. Proust, *Los placeres y los días: Parodias y miscelánea*. Alianza, 1975. 9

[3] H. K. Khalil, *Noninear systems*, ., Ed. Prentice-Hall, 1996, vol. 2, no. 5. 1, 42, 46

[4] M. W. Spong, S. Hutchinson, and V. M., *Robot Modeling and Control*, ., Ed. John Wiley and Sons, Inc., 2006, vol. 141, no. 1. 1, 24

[5] A. Isidori, *Nonlinear Control Systems*. Springer Science & Business Media, 2013. 1

[6] C. Chen, *Linear System Theory and Design*. Oxford University Press, Inc., 1998. 1

[7] D. S. Naidu, *Optimal Control Systems*. CRC Press, 2002. 2, 3, 5, 7, 8, 16

[8] H. J. Sussmann and J. C. Willems, "300 Years of Optimal Control: from the Brachystochrone to the Maximum Principle," *IEEE Control Systems*, vol. 17, no. 3, pp. 32–44, 1997. 2

[9] D. E. Kirk, *Optimal Control Theory: An Introduction*, ser. Dover Books on Electrical Engineering Series. Dover Publications, 2004. 2, 5, 7, 8

[10] R. Bellman, *Dynamic Programming*. Princeton University Press, 1957. 2, 6, 7, 16, 26

[11] A. E. Bryson, *Applied Optimal Control: Optimization, Estimation and Control*. CRC Press, 1975. 2, 7, 8, 16

[12] D. Liberzon, *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton University Press, 2012. 2, 7, 8, 24, 36

[13] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific Belmont, MA, 1995, vol. 1, no. 2. 2, 7, 16, 26

[14] M. G. Crandall and P. L. Lions, "Viscosity solutions of Hamilton-Jacobi equations in infinite dimensions. IV. Hamiltonians with unbounded linear terms," *Journal of Functional Analysis*, vol. 90, no. 2, pp. 237 – 283, 1990. 3

[15] R. W. Beard, G. N. Saridis, and J. T. Wen, "Approximate Solutions to the Time-Invariant Hamilton-Jacobi-Bellman Equation," *Journal of Optimization Theory and Applications*, vol. 96, no. 3, pp. 589–626, Mar 1998. 3, 14, 17

[16] J. H. Lee, "Model Predictive Control and Dynamic Programming," in *Control, Automation and Systems (ICCAS), 2011 11th International Conference on*. IEEE, 2011, pp. 1807–1809. 3

[17] F. L. Lewis and D. Vrabie, "Reinforcement Learning and Adaptive Dynamic Programming for Feedback Control," *IEEE Circuits and Systems Magazine*, vol. 9, no. 3, pp. 32–50, 2009. 3, 15, 18

[18] N. K. Masmoudi, C. Rekik, M. Djemel, and N. Derbel, "Two coupled neural-networks-based solution of the Hamilton–Jacobi–Bellman equation," *Applied Soft Computing*, vol. 11, no. 3, pp. 2946 – 2963, 2011. 3, 18

[19] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359 – 366, 1989. 3, 12, 14, 18

[20] L. Fridman, A. Poznyak, F. J. Bejarano, *et al.*, *Robust output LQ optimal control via integral sliding modes*. Springer, 2014. 3

[21] S. D. Mulje and R.M.Nagarele, "LQR Technique based Second Order Sliding Mode Control for Linear Uncertain Systems," *International Journal of Computer Applications*, vol. 137, no. 7, pp. 23–29, March 2016. 3, 17

[22] A. S. Poznyak and V. G. Boltyanski, *The Robust Maximum Principle*, T. Bašar, Ed. Springer Science & Business Media, 2012. 3, 7, 16, 17, 18, 19, 25

[23] D. P. Bertsekas, "Neuro-Dynamic Programming," in *Encyclopedia of Optimization*. Springer, 2008, pp. 2555–2560. 4

[24] L. S. Pontryagin, *Mathematical Theory of Optimal Processes*. Routledge, 2018. 5

[25] R. E. Bellman and S. E. Dreyfus, *Applied Dynamic Programming*. Princeton University Press, 2015, vol. 2050. 5

[26] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control*. John Wiley & Sons, 2012. 7, 8

[27] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Prentice Hall PTR, 1994. 12, 14, 50

[28] S. Jagannathan and F. L. Lewis, "Identification of nonlinear dynamical systems using multilayered neural networks," *Automatica*, vol. 32, no. 12, pp. 1707–1712, 1996. 12, 50

[29] F. W. Lewis, S. Jagannathan, and A. Yesildirak, *Neural Network Control of Robot Manipulators and Non-linear Systems*. CRC Press, 1998. 12, 16, 50

[30] A. S. Poznyak, E. N. Sanchez, and W. Yu, *Differential Neural Networks for Robust Nonlinear Control*, B. L. C. in Publication Data, Ed. World Scientific Publishing, 2001. 12, 14, 18, 50, 55, 78, 79

[31] A. R. Barron, "Universal approximation bounds for superpositions of a sigmoidal function," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 930–945, 1993. 12, 14, 15

[32] J. C. Patra, R. N. Pal, B. Chatterji, and G. Panda, "Identification of nonlinear dynamic systems using functional link artificial neural networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 29, no. 2, pp. 254–262, 1999. 12

[33] Y. Shirvany, M. Hayati, and R. Moradian, "Multilayer perceptron neural networks with novel unsupervised training method for numerical solution of the partial differential equations," *Applied Soft Computing*, vol. 9, no. 1, pp. 20–29, 2009. 13

[34] M. Kumar and N. Yadav, "Multilayer perceptrons and radial basis function neural network methods for the solution of differential equations: a survey," *Computers & Mathematics with Applications*, vol. 62, no. 10, pp. 3796–3811, 2011. 13

[35] P. Sibi, S. A. Jones, and P. Siddarth, "Analysis of different activation functions using back propagation neural networks," *Journal of Theoretical and Applied Information Technology*, vol. 47, no. 3, pp. 1264–1268, 2013. 14

[36] A. Mohemmed, S. Schliebs, S. Matsuda, and N. Kasabov, "Span: Spike pattern association neuron for learning spatio-temporal spike patterns," *International Journal of Neural Systems*, vol. 22, no. 04, p. 1250012, 2012. 14

[37] M. Alfaro-Ponce, A. Argüelles, I. Chairez, and A. Pérez, "Automatic electroencephalographic information classifier based on recurrent neural networks," *International Journal of Machine Learning and Cybernetics*, pp. 1–13, 2018. 14, 18

[38] Q. Xue, Y. H. Hu, and W. J. Tompkins, "Neural-network-based adaptive matched filtering for QRS detection," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 4, pp. 317–329, 1992. 14

[39] L. Yu, S. Wang, and K. K. Lai, "Adaptive smoothing neural networks in foreign exchange rate forecasting," in *International Conference on Computational Science*. Springer, 2005, pp. 523–530. 14

[40] A. Bhattacharya and C. Chakraborty, "A shunt active power filter with enhanced performance using ANN-based predictive and adaptive controllers," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 2, pp. 421–428, 2010. 14

[41] A. N. Lakhal, A. S. Tlili, and N. B. Braiek, "Neural network observer for nonlinear systems application to induction motors," *International Journal of Control and Automation*, vol. 3, no. 1, pp. 1–16, 2010. 14

[42] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," *IEEE Transactions on Neural Networks*, vol. 1, no. 1, pp. 4–27, 1990. 14, 18

[43] A. H. E. Zooghby, C. G. Christodoulou, and M. Georgiopoulos, "Neural network-based adaptive beamforming for one-and two-dimensional antenna arrays," *IEEE Transactions on Antennas and Propagation*, vol. 46, no. 12, pp. 1891–1893, 1998. 14

[44] J. A. Leonard, M. A. Kramer, and L. H. Ungar, "Using radial basis functions to approximate a function and its error bounds," *IEEE Transactions on Neural Networks*, vol. 3, no. 4, pp. 624–627, 1992. 14

[45] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of control, signals and systems*, vol. 2, no. 4, pp. 303–314, 1989. 14, 15, 50

[46] J. Park and I. W. Sandberg, "Universal approximation using radial-basis-function networks," *Neural Computation*, vol. 3, no. 2, pp. 246–257, 1991. 14, 18

[47] R. Hecht-Nielsen, "Kolmogorov's mapping Neural Network existence Theorem," in *Proceedings of the International Conference on Neural Networks*, vol. 3. IEEE Press New York, 1987, pp. 11–14. 14

[48] K. Hornik, M. Stinchcombe, and H. White, "Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks," *Neural Networks*, vol. 3, no. 5, pp. 551 – 560, 1990. 14

[49] V. Kurková, "Kolmogorov's theorem and multilayer neural networks," *Neural Networks*, vol. 5, no. 3, pp. 501–506, 1992. 14

[50] N. E. Cotter, "The Stone-Weierstrass theorem and its application to neural networks," *IEEE Transactions on Neural Network*, vol. 1, pp. 290–295, 1990. 14, 52

[51] K.-I. Funahashi, "On the approximate realization of continuous mappings by neural networks," *Neural Networks*, vol. 2, no. 3, pp. 183–192, 1989. 14, 15

[52] P. Koiran, "On the complexity of approximating mappings using feedforward networks," *Neural Networks*, vol. 6, no. 5, pp. 649–653, 1993. 14

[53] P. C. Kainen, V. Kurkova, and M. Sanguineti, "Dependence of computational models on input dimension: Tractability of approximation and optimization tasks," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1203–1214, 2012. 14

[54] N. J. Guliyev and V. E. Ismailov, "On the approximation by single hidden layer feedforward neural networks with fixed weights," *Neural Networks*, vol. 98, pp. 296–304, 2018. 14

[55] D. Vrabie and F. Lewis, "Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems," *Neural Networks*, vol. 22, no. 3, pp. 237 – 246, 2009. 14, 17, 18, 19

[56] D. Wang, D. Liu, H. Li, and H. Ma, "Neural-network-based robust optimal control design for a class of uncertain nonlinear systems via adaptive dynamic programming," *Information Sciences*, vol. 282, pp. 167–179, 2014. 14, 17

[57] T. Cheng, F. L. Lewis, and M. Abu-Khalaf, "A neural network solution for fixed-final time optimal control of nonlinear systems," *Automatica*, vol. 43, no. 3, pp. 482 – 490, 2007. 14, 18

[58] K. G. Vamvoudakis and F. L. Lewis, "Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878 – 888, 2010. 14, 18

[59] A. Heydari and S. N. Balakrishnan, "Finite-Horizon Control-Constrained Nonlinear Optimal Control Using Single Network Adaptive Critics," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 1, pp. 145–157, Jan 2013. 14, 18

[60] B. Igelnik and Y.-H. Pao, "Stochastic choice of basis functions in adaptive function approximation and the functional-link net," *IEEE Transactions on Neural Networks*, vol. 6, no. 6, pp. 1320–1329, Nov 1995. 16

[61] L. S. Pontryagin, *Mathematical theory of optimal processes*. Routledge, 1987. 16

[62] M. Palanisamy, H. Modares, F. L. Lewis, and M. Aurangzeb, "Continuous-Time Q-Learning for Infinite-Horizon Discounted Cost Linear Quadratic Regulator Problems," *IEEE Transactions on Cybernetics*, vol. 45, no. 2, pp. 165–176, Feb 2015. 16

[63] A. P. Sage, *Optimum Systems Control*. Prentice-Hall, 1968. 16, 36

[64] V. Utkin, J. Guldner, and J. Shi, *Sliding Mode Control in Electro-Mechanical Systems*. CRC Press, 2009, vol. 34. 16

[65] C. Edwards and S. Spurgeon, *Sliding Mode Control: Theory and Applications*. CRC Press, 1998. 16

[66] A. Poznyak, A. Polyakov, and V. Azhmyakov, *Attractive Ellipsoids in Robust Control*. Springer, 2014. 16, 43

[67] V. Azhmyakov, "On the geometric aspects of the invariant ellipsoid method: Application to the robust control design," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, Dec 2011, pp. 1353–1358. 16

[68] Q. Yang, S. Jagannathan, and Y. Sun, "Robust integral of neural network and error sign control of mimo nonlinear systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 12, pp. 3278–3286, Dec 2015. 16

[69] Z. L. Tang, S. S. Ge, K. P. Tee, and W. He, "Robust adaptive neural tracking control for a class of perturbed uncertain nonlinear systems with state constraints," *IEEE Transactions*

on Systems, Man, and Cybernetics: Systems, vol. 46, no. 12, pp. 1618–1629, Dec 2016. 16

[70] J. D. Pearson, "Approximation methods in optimal control I. Sub-optimal control," *International Journal of Electronics*, vol. 13, no. 5, pp. 453–469, 1962. 17

[71] H. Modares, F. L. Lewis, and M.-B. Naghibi-Sistani, "Integral reinforcement learning and experience replay for adaptive optimal control of partially-unknown constrained-input continuous-time systems," *Automatica*, vol. 50, no. 1, pp. 193 – 202, 2014. 17, 18, 19

[72] R. W. Beard, G. N. Saridis, and J. T. Wen, "Galerkin approximations of the generalized Hamilton-Jacobi-Bellman equation," *Automatica*, vol. 33, no. 12, pp. 2159 – 2177, 1997. 17

[73] I. C. Dolcetta, "On a discrete approximation of the Hamilton-Jacobi equation of Dynamic Programming," *Applied Mathematics and Optimization*, vol. 10, no. 1, pp. 367–377, 1983. 17

[74] L. Grüne, "An adaptive grid scheme for the discrete Hamilton-Jacobi-Bellman equation," *Numerische Mathematik*, vol. 75, no. 3, pp. 319–337, 1997. 17

[75] J. R. Cloutier and J. C. Cockburn, "The state-dependent nonlinear regulator with state constrains," in *Proceedings*, 2001. 17

[76] A. Heydari and S. N. Balakrishnan, "Approximate closed-form solutions to finite-horizon optimal control of nonlinear systems," in *2012 American Control Conference (ACC)*, June 2012, pp. 2657–2662. 17

[77] J. R. Cloutier, "State-dependent Riccati equation techniques: an overview," in *Proceedings of the 1997 American Control Conference (Cat. No. 97CH36041)*, vol. 2. IEEE, 1997, pp. 932–936. 17

[78] T. Çimen, "State-Dependent Riccati Equation (SDRE) Control: A Survey," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 3761 – 3775, 2008, 17th IFAC World Congress. 17

[79] C. S. Huang, S. Wang, and K. L. Teo, "Solving Hamilton-Jacobi-Bellman equations by a modified method of characteristics," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 40, no. 1, pp. 279 – 293, 2000. 17

[80] S. Wang, L. S. Jennings, and K. L. Teo, "Numerical Solution of Hamilton-Jacobi-Bellman Equations by an Upwind Finite Volume Method," *Journal of Global Optimization*, vol. 27, no. 2, pp. 177–192, Nov 2003. 17

[81] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Springer Science & Business Media, 2008. 17

[82] H. S. Nik, S. Effati, and M. Shirazian, "An approximate-analytical solution for the Hamilton–Jacobi–Bellman equation via homotopy perturbation method," *Applied Mathematical Modelling*, vol. 36, no. 11, pp. 5614–5623, 2012. 18

[83] Y. H. Kim, F. L. Lewis, and D. M. Dawson, "Intelligent optimal control of robotic manipulators using neural networks," *Automatica*, vol. 36, no. 9, pp. 1355 – 1364, 2000. 18

[84] S. Ferrari and R. F. Stengel, "Smooth function approximation using neural networks," *IEEE Transactions on Neural Networks*, vol. 16, no. 1, pp. 24–38, 2005. 18

[85] R. S. Beidokhti and A. Malek, "Solving initial-boundary value problems for systems of partial differential equations using neural networks and optimization techniques," *Journal of the Franklin Institute*, vol. 346, no. 9, pp. 898–913, 2009. 18

[86] M. Abu-Khalaf and F. L. Lewis, "Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach," *Automatica*, vol. 41, no. 5, pp. 779–791, 2005. 18

[87] K. G. Vamvoudakis, D. Vrabie, and F. L. Lewis, "Online adaptive algorithm for optimal control with integral reinforcement learning," *International Journal of Robust and Nonlinear Control*, vol. 24, no. 17, pp. 2686–2710, 2014. 18

[88] B. Kiumarsi, F. L. Lewis, H. Modares, A. Karimpour, and M.-B. Naghibi-Sistani, "Reinforcement Q-learning for optimal tracking control of linear discrete-time systems with unknown dynamics," *Automatica*, vol. 50, no. 4, pp. 1167 – 1175, 2014. 18

[89] R. S. Sutton and A. G. Barto, *Reinforcement Learning an Introduction*, 2nd ed. The MIT Press, 2012. 18

[90] Y. Jiang and Z. P. Jiang, "Robust adaptive dynamic programming and feedback stabilization of nonlinear systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 5, pp. 882–893, May 2014. 18

[91] D. Liu, X. Yang, and H. Li, "Adaptive optimal control for a class of continuous-time affine nonlinear systems with unknown internal dynamics," *Neural Computing and Applications*, vol. 23, no. 7, pp. 1843–1850, Dec 2013. 19

[92] H. Modares, F. L. Lewis, and Z.-P. Jiang, "Optimal output-feedback control of unknown continuous-time linear systems using off-policy reinforcement learning," *IEEE Transactions on Cybernetics*, vol. 46, no. 11, pp. 2401–2410, 2016. 19

[93] H. Hermes, "Nilpotent approximations of control systems and distributions," *SIAM Journal on Control and Optimization*, vol. 24, no. 4, pp. 731–736, 1986. 20, 50

[94] V. Andrieu, L. Praly, and A. Astolfi, "Homogeneous approximation, recursive observer design, and output feedback," *SIAM Journal on Control and Optimization*, vol. 47, no. 4, pp. 1814–1850, 2008. 20, 50

[95] Y. Orlov, "Finite time stability and robust control synthesis of uncertain switched systems," *SIAM Journal on Control and Optimization*, vol. 43, no. 4, pp. 1253–1271, 2005. 20, 50

[96] A. Levant, "Homogeneity approach to high-order sliding mode design," *Automatica*, vol. 41, no. 5, pp. 823–830, 2005. 20, 50

[97] K. Zimenko, D. Efimov, A. Polyakov, and W. Perruquetti, "A note on delay robustness for homogeneous systems with negative degree," *Automatica*, vol. 79, pp. 178–184, 2017. 20, 50, 76

[98] J.-B. Pomet and C. Samson, "Time-varying exponential stabilization of nonholonomic systems in power form," Ph.D. dissertation, Inria, 1993. 20, 50

[99] H. Hermes, "Homogeneous feedback controls for homogeneous systems," *Systems & Control Letters*, vol. 24, no. 1, pp. 7 – 11, 1995. 20, 50

[100] E. Bernuau, D. Efimov, W. Perruquetti, and A. Polyakov, "On homogeneity and its application in sliding mode control," *Journal of the Franklin Institute*, vol. 351, no. 4, pp. 1866–1901, 2014, Special Issue on 2010-2012 Advances in Variable Structure Systems and Sliding Mode Algorithms. 20, 50

[101] A. Polyakov, D. Efimov, E. Fridman, and W. Perruquetti, "On Homogeneous Distributed Parameter Systems," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3657–3662, Nov 2016. 20, 50, 51

[102] S. Nazari and B. Shafai, "Robust SDC parameterization for a class of extended linearization systems," in *Proceedings of the 2011 American Control Conference*. IEEE, 2011, pp. 3742–3747. 24

[103] B. Friedland, *Advanced Control System Design*. Prentice-Hall, Inc., 1995. 24

[104] F. R. Gantmacher and J. L. Brenner, *Applications of the Theory of Matrices*. Courier Corporation, 2005. 29

[105] D. S. Bernstein, *Scalar, Vector, and Matrix Mathematics: Theory, Facts, and Formulas-Revised and Expanded Edition*. Princeton University Press, 2018. 29, 30, 46, 57

[106] B. D. O. Anderson and J. B. Moore, *Optimal control: Linear Quadratic Methods*. Courier Corporation, 2007. 36

[107] R. Behling, D. S. Gonçalves, and S. A. Santos, "Local Convergence Analysis of the Levenberg–Marquardt Framework for Nonzero-Residue Nonlinear Least-Squares Problems Under an Error Bound Condition," *Journal of Optimization Theory and Applications*, vol. 183, no. 3, pp. 1099–1122, 2019. 39

[108] E. Bergou, Y. Diouane, and V. Kungurtsev, "Global and Local Convergence of a Levenberg-Marquadt Algorithm for Inverse Problems," Technical Report ISAE-SUPAERO, Tech. Rep., 2017. 39

[109] W. M. Haddad and V. Chellaboina, *Nonlinear dynamical systems and control: a Lyapunov-based approach*, ., Ed. Princeton University Press, 2011. 42

[110] L. Ljung, "Some aspects on Nonlinear System Identification," *IFAC Proceedings Volumes*, vol. 39, no. 1, pp. 553 – 564, 2006, 14th IFAC Symposium on Identification and System Parameter Estimation. 49

[111] S. A. Billings, "Identification of nonlinear systems-a survey," in *IEEE Proceedings D-Control Theory and Applications*, vol. 127, no. 6. IET, 1980, pp. 272–285. 50

[112] R. Haber and L. Keviczky, *Nonlinear System Identification Input-Output Modeling Approach*. Kluwer Academic Publishers, 1999. 50

[113] O. Nelles, *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*. Springer Science & Business Media, 2013. 50

[114] E. D. Sontag, "Some topics in neural networks and control," in *Proceedings of the European Control Conference*, 1993. 50

[115] P. Diaconis and M. Shahshahani, "On nonlinear functions of linear combinations," *SIAM Journal on Scientific and Statistical Computing*, vol. 5, no. 1, pp. 175–191, 1984. 50

[116] S. Hubbert, "Radial basis function interpolation on the sphere," Ph.D. dissertation, University of London, 2002. 50

[117] I. Chairez, "Adaptive neural network nonparametric identifier with normalized learning laws," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 5, pp. 1216–1227, 2017. 50, 55

[118] V. Zubov, "On systems of ordinary differential equations with generalized homogenous right-hand sides," *Izvestiya Vuzov, Matematika*, vol. 1, no. 2, pp. 80–88, 1958. 51

[119] V. V. Khomenuk, "(in russian)," *Izvestiya Vuzov, Matematika*, vol. 3, no. 22, pp. 157–164, 1961. 51

[120] M. Kawski, "Geometric homogeneity and stabilization," in *Nonlinear Control Systems Design 1995*. Elsevier, 1995, pp. 147–152. 51

[121] L. Grüne, "Homogeneous state feedback stabilization of homogenous systems," *SIAM Journal on Control and Optimization*, vol. 38, no. 4, pp. 1288–1308, 2000. 51

[122] A. Polyakov, "Sliding mode control design using canonical homogeneous norm," *International Journal of Robust and Nonlinear Control*, vol. 0, no. 0, 2018. 51, 53

[123] S. P. Bhat and D. S. Bernstein, "Geometric homogeneity with applications to finite-time stability," *Mathematics of Control, Signals and Systems*, vol. 17, no. 2, pp. 101–127, 2005. 53

[124] D. Efimov and A. Fradkov, "Design of impulsive adaptive observers for improvement of persistency of excitation," *International Journal of Adaptive Control and Signal Processing*, vol. 29, no. 6, pp. 765–782, 2015. 55, 104, 105

[125] C. Join, H. Sira-Ramírez, and M. Fliess, "Control of an uncertain three-tank system via on-line parameter indentification and fault detection," *IFAC Proceedings Volumes*, vol. 38, no. 1, pp. 251 – 256, 2005, 16th IFAC World Congress. 76

[126] R. Seydou, T. Raissi, A. Zolghadri, and D. Efimov, "Actuator fault diagnosis for flat systems: a constraint satisfaction approach," *International Journal of Applied Mathematics and Computer Science*, vol. 23, no. 1, pp. 171–181, 2013. 76

Appendices

Appendix A

Complementary Results

A.1 Proof of boundedness for the auxiliary variables

This appendix presents the lemmas to prove the boundedness for δ and Ω_i.

Lemma 1. *Consider the following dynamical system:*

$$\dot{r} = \varphi_1(t) A r + \varphi_2(t), \tag{A.1}$$

where $r \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$ is a Hurwitz matrix, $\varphi_1 : \mathbb{R}_+ \to \mathbb{R}_+$ is locally bounded separated from zero

$$\varphi_1(t) \geq \beta_1 > 0, \quad \forall t \in \mathbb{R}_+, \tag{A.2}$$

and the function $\varphi_2 : \mathbb{R}_+ \to \mathbb{R}^n$ is globally bounded $\sup_{t\in\mathbb{R}_+} \|\varphi_2(t)\| = \varphi_2^+ < +\infty$. Then, there exists a class-$\mathcal{K}$ function $\rho_r(\cdot)$ such that the solutions of (A.1) *satisfy*

$$\limsup_{t\to\infty} \|r(t)\| \leq \rho_r\left(\varphi_2^+\right), \tag{A.3}$$

Proof. Consider the Lyapunov function candidate $V_r(r) = r^\top P r$, the time derivative is:

$$\dot{V}_r = \varphi_1 r^\top \left(A^\top P + PA\right) r + 2\varphi_2^\top P r, \tag{A.4}$$

where $P \in \mathbb{R}^{n\times n}$, $P = P^\top > 0$, such that:

$$A^\top P + PA \leq -\gamma P,\ \gamma \in \mathbb{R}_+, \tag{A.5}$$

Using the Λ-inequality we derive

$$2\varphi_2^\top Pr \leq \|\varphi_2\|_\Lambda^2 + \|Pr\|_{\Lambda^{-1}}^2, \tag{A.6}$$

for any $\Lambda \in \mathbb{R}^{n\times n}$, $\Lambda = \Lambda^\top$, $\Lambda > 0$. By taking $\Lambda = 2\beta_1^{-1}\gamma^{-1}P$ in (A.6), considering (A.5) and the condition (A.2), the following relation holds:

$$\dot{V}_r \leq -0.5\beta_1\gamma V_r(r) + 2\beta_1^{-1}\gamma^{-1}\lambda_{\max}\{P\}\,\|\varphi_2(t)\|^2, \tag{A.7}$$

Hence, we derive

$$\limsup_{t\to+\infty} V_r(t) \leq \frac{2\beta_1^{-1}\gamma^{-1}\lambda_{\max}\{P\}\,(\varphi_2^+)^2}{0.5\beta_1\gamma}$$

and (A.3) is fulfilled. □

A.2 Persistent excitation condition

The following result is a Corollary of [124, Lemma 1].

Corollary 3. *Consider the time-varying dynamical system:*

$$\dot{z}(t) = -\gamma_z(t)K_z G_z^\top(t)G_z(t)z(t) + v(t),\ t \in \mathbb{R}_+, \tag{A.8}$$

where $z(t) \in \mathbb{R}^k$ is the state of the system, the continuous function $\gamma_z : \mathbb{R}_+ \to \mathbb{R}_+$ is positive and bounded $0 < \gamma^- \leq \gamma_z(t) \leq \gamma^+$, $\forall t \in \mathbb{R}_+$, the continuous vector-valued function $v : \mathbb{R}_+ \to \mathbb{R}^k$ is uniformly bounded and $\limsup_{t\to+\infty}\|v(t)\| \leq v^+$.

The symmetric matrix $K_z \in \mathbb{R}^{k\times k}$ is positive definite and the continuous matrix-valued function $G_z : \mathbb{R}_+ \to \mathbb{R}^{q\times k}$ is uniformly bounded. If the following PE condition

$$\int_t^{t+\ell} G_z^\top(s)G_z(s)ds \geq \vartheta I_k, \quad \forall t \in \mathbb{R}_+ \tag{A.9}$$

holds for some $\vartheta > 0$ and $\ell > 0$, then, there exists $\rho_z \in \mathcal{K}$ such that for any initial condition $z(0) \in \mathbb{R}^k$, the solution $z(t)$ of the system (A.8) *is defined for all $t \in \mathbb{R}_+$ and $\limsup_{t\to+\infty}\|z(t)\| \leq \rho_z(v^+)$.*

Proof. Applying the following change of variable $\tilde{z} = K_z^{-1/2} z$, we derive

$$\frac{d\tilde{z}(t)}{dt} = -\gamma_z(t) K_z^{1/2} G_z^\top(s) G_z(s) K_z^{1/2} \tilde{z}(t) + K_z^{-1/2} v(t)$$

Then,

$$-\gamma_z(t) K_z^{1/2} G_z^\top(s) G_z(s) K_z^{1/2} \tilde{z}(t) + K_z^{-1/2} v(t) = -\tilde{G}^\top(s) \tilde{G}(s) \tilde{z}(t) + \tilde{v}(t),$$

where $\tilde{G}(t) = \sqrt{\gamma_z(t)} G_z(t) K_z^{1/2}$ and $\tilde{v}(t) = K_z^{-1/2} v(t)$. Obviously, if the PE condition holds, then

$$\int_t^{t+\ell} \tilde{G}^\top(s) \tilde{G}(s) ds = K_z^{1/2} \left(\int_t^{t+\ell} \gamma_z(s) G_z^\top(s) G_z(s) ds \right) K_z^{1/2}.$$

Considering γ^- the following inequality holds

$$\int_t^{t+\ell} \tilde{G}^\top(s) \tilde{G}(s) ds \geq \gamma^- K_z^{1/2} \left(\int_t^{t+\ell} G_z^\top(s) G_z(s) ds \right) K_z^{1/2}.$$

Under (A.9),

$$\int_t^{t+\ell} \tilde{G}^\top(s) \tilde{G}(s) ds \geq \gamma^- K_z^{1/2} (\vartheta I_k) K_z^{1/2},$$

Therefore, as $\gamma^- K_z^{1/2} (\vartheta I_k) K_z^{1/2} = \gamma^- \vartheta K_z$, we obtain

$$\int_t^{t+\ell} \tilde{G}^\top(s) \tilde{G}(s) ds \geq \gamma^- \vartheta \lambda_{\min}(K_z) I_k, \quad \forall t \in \mathbb{R}_+.$$

Finally, applying [124, Lemma 1] we complete the proof. □

Appendix B

Publications

B.1 Conferences

1. **11th IFAC Symposium on Nonlinear Control Systems NOLCOS 2019 Vienna, Austria, 4–6 September 2019.**

   Ballesteros, M., Polyakov, A., Efimov, D., Chairez, I. and Poznyak, A. (2019). Differential Neural Network Identification for Homogeneous Dynamical Systems. IFAC-PapersOnLine, 52(16), 233-238. DOI: 10.1016/j.ifacol.2019.11.784

2. **2nd IFAC Conference on Modelling, Identification and Control of Nonlinear Systems MICNON 2018, Guadalajara, Jalisco, Mexico, 20–22 June 2018.**

   Ballesteros-Escamilla, M., Chairez, I., Boltyanski, V. G., and Poznyak, A. (2018). Realization of robust optimal control by dynamic neural-programming. IFAC-PapersOnLine, 51(13), 468-473. DOI:10.1016/j.ifacol.2018.07.322.

B.2 Journal Articles

1. Ballesteros, M., Chairez, I. and Poznyak, A. (2020). Robust Min-Max Optimal Control Design for Systems with Uncertain Models: a Neural Dynamic Programming Approach. Neural Networks. DOI: 10.1016/j.neunet.2020.01.016.

B.3 Submitted Articles

1. Mariana Ballesteros, Alexander Poznyak and Isaac Chairez. Robust optimal feedback control design for uncertain systems based on artificial neural network approximation of the Bellman's value function. IEEE Transactions on Neural Networks and Learning Systems.

2. Mariana Ballesteros-Escamilla, Andrey Polyakov, Denis Efimov, Isaac Chairez and Alexander S. Poznyak. Non-parametric Identification of Homogeneous Dynamical Systems. Automatica.

3. 21st IFAC World Congress in Berlin, Germany, July 12-17, 2020.

   Mariana Ballesteros-Escamilla, Andrey Polyakov, Denis Efimov, Isaac Chairez and Alexander S. Poznyak. Adaptive Discontinuous Control for Homogeneous Systems Approximated by Neural Networks.