# Non-parametric Identification of Homogeneous Dynamical Systems

Mariana Ballesteros, Andrey Polyakov, Denis Efimov, Isaac Chairez,
Alexander Poznyak

**HAL Id: hal-03141177**

**https://hal.inria.fr/hal-03141177**

Submitted on 15 Feb 2021

# Non-parametric Identification of Homogeneous Dynamical Systems [*]

Mariana Ballesteros [a,b], Andrey Polyakov [b,c], Denis Efimov [b,c],
Isaac Chairez [d], Alexander Poznyak [a]

[a]*Department of Automatic Control, CINVESTAV-IPN Av. IPN 2508, San Pedro Zacatenco 07360, Mexico City*
[b]*Inria, Univ. Lille, CNRS, UMR 9189 - CRIStAL, F-59000 Lille, France*
[c]*ITMO University, 49 Av. Kronverkskiy, 197101 Saint Petersburg, Russia*
[d]*Department of Bioprocesses, Unidad Profesional Interdisciplinaria de Biotecnología-Instituto Politécnico Nacional, Mexico City*

## Abstract

The aim of this study is to design a non-parametric identifier for homogeneous systems based on a class of artificial neural networks with continuous dynamics. The identification algorithm is developed for input-affine systems with uncertain gains and diverse degrees of homogeneity. One of the main contributions of this study is the extension of the universal approximation property of neural networks for continuous homogeneous systems. Another contribution is the development of a differential non-parametric identifier based on the novel concept of homogeneous neural networks. The adjustment laws for the weights are obtained from a Lyapunov stability analysis taking homogeneity properties of the system into account. The ultimate boundedness of the origin for the identification error is demonstrated using the persistent excitation condition. The effectiveness of the proposed identifier is verified by the simulation of the three-tank homogeneous model. In this example, the proposed identification scheme is compared with a classical ANN identifier, and we present a statistical analysis of such comparison. It is shown in simulations that the identification error of the proposed homogeneous algorithm has faster convergence and less oscillations.

*Key words:* Non-Parametric Identification, Homogeneous Systems, Differential Neural Networks, Approximate Homogeneity.

## 1 Introduction

Control and optimization of uncertain systems is an active research area in the modern control theory. One option to control such a class of systems consists of designing compensating controllers based on an approximated model of the uncertain dynamics. This modeling technique is usually known as system identification [1]. The identifier design can be either parametric or non-parametric. The first one assumes the existence of an adequate mathematical representation but with uncertain parameters, which must be estimated using the input-output response. The second approach proposes an approximate mathematical model whose parameters must be estimated [2,3]. The identification method for the adjustment of the parameters must guarantee at least that the identification error is ultimately bounded.

The non-parametric identification problem has been tackled using several methods due to the inherent complexity of uncertain models of real plants. Some of these identification techniques are also known as black-box identification using functional series methods, nonlinear auto-regressive moving average with exogenous input, frequency domain approaches, fuzzy models, and artificial neural networks (ANNs) [2,4–6]. Differential neural networks (DNNs), which form a class of ANNs with a continuous evolution

of its states, have been used to solve the non-parametric identification problem of dynamical systems [7,8]. In [9], the identification, tracking control and disturbance rejection for a class of uncertain system is proposed using DNNs. This variant of ANN offers a mathematical structure to approximate a large class of nonlinear time-dependent mappings, see [10,11]. DNNs can be adjusted on-line and used for multi-input-multi-output systems. The class of DNN is defined by the activation functions, the number of layers, and the kind of interconnection between them [12,13]. For example in [14] and [15] the network structure uses radial basis functions for the identification task. Besides the DNN structure, the development of a DNN identifier needs the design of the parameter adjustment algorithms. There are diverse methods to adapt the parameters based on continuous versions of the least mean square technique or the Lyapunov stability theory for systems with time-varying parameters [16]. In [17], the parameters are adjusted using a Lyapunov stability analysis but for a recurrent neural network, in the discrete-time case. Most of these techniques offer similar efficient approximations. However, an additional a priori information about homogeneity (dilation symmetry) of an uncertain system may allow better identification of the system. The lack of a solution using symmetry information is the motivation of developing non-parametric identifiers with continuous dynamics for homogeneous systems.

Homogeneity is a kind of symmetry under which an object remains consistent for a certain scaling or dilation. Homogeneous systems can be utilized for local approximations [18,19] or set-valued extensions [20,21] of nonlinear control

systems. In particular, some models of control process [22], non-holonomic mechanical systems [23], and systems with frictions [20] 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 [24–27].

To the best of the authors' knowledge, the identification problem of homogeneous systems has not been profoundly studied in the literature. The need in such an identifier is motivated by the wide class of uncertain systems which preserve the homogeneity property (at least locally). Based on all the above mentioned homogeneity properties and the wide application of the non-parametric identifiers for controlling uncertain systems, this study offers the design of a non-parametric identifier with a homogeneous structure using DNNs. In consequence, if the identifier may approximate locally a homogeneous system, then the approximation ability can be expanded to the whole state space [24,26,27].

A preliminary version of this work was presented in [28], the **novelty** of this work consists of the following:

-The key difference of the present version is that the stability analysis of the DNN identifier in the case of a non-exact representation of the system is provided; the proofs and the comparison of the presented homogeneous DNN identifier with a classical one are given.
-For homogeneous and locally homogeneous dynamical systems, a non-parametric identifier is designed based on DNN.
-The universal approximation property of ANN for homogeneous systems is quantified.
-The closed-loop identification error stability is proven using Lyapunov arguments, homogeneity and persistence of excitation (PE) condition.
-Numerical validation of the proposed DNN-identifier for a three-tank system confirms efficiency of the proposed method.

*Notation:* $\mathbb{R}$ denotes the set of real numbers, $\mathbb{R}_+ = \{x \in \mathbb{R}: x \geq 0\}$. For any $\vartheta \in \mathbb{R}_+$ and $\forall x \in \mathbb{R}$ we set $\lceil x \rfloor^\vartheta = \text{sign}(x)|x|^\vartheta$. The Euclidean norm is denoted by $\| \cdot \|$, the weighted norm $\|z\|_Q := \sqrt{z^\top Q z}$ for $z \in \mathbb{R}^n$ and $Q = Q^\top \in \mathbb{R}^{n \times n}$, $Q > 0$. The trace of matrix $H \in \mathbb{R}^{n \times n}$ is defined as $\text{tr}\{H\}$. The Frobenius norm is denoted by $\|B\|_F = \sqrt{\text{tr}\{B^\top B\}}$ for $B \in \mathbb{R}^{n \times r}$. We set $\text{vec}(M) \in \mathbb{R}^{nm}$ as the vectorization of matrix $M \in \mathbb{R}^{n \times m}$. The Kronecker product is denoted by $\otimes$. Notation $\text{diag}\{A_1, \ldots, A_n\}$ means the block diagonal matrix conformed by matrices $A_i$, $i = \{1, \ldots, n\}$. A continuous function $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ is a class-$\mathcal{K}$ function if it is strictly increasing and $\rho(0) = 0$.

## 2 Problem Statement

The nonlinear system considered in this manuscript is given by the following ordinary differential equations:

$$\dot{x} = f(x, u), \quad t \in \mathbb{R}_+, \quad x(0) = x_0, \qquad (1)$$

where $f(x, u) := f_0(x) + \sum_{i=1}^m f_i(x)u_i$, $x(t) \in \mathbb{R}^n$ is the state vector of the system; $x_0 \in \mathbb{R}^n$ is the initial condition, $u(t) = [u_1(t), .., u_m(t)]^\top \in \mathbb{R}^m$ is the control input, $m \leq n$. The function $f_0 : \mathbb{R}^n \to \mathbb{R}^n$ is a drift term of the system, the input associated functions $f_i : \mathbb{R}^n \to \mathbb{R}^n$, $i = 0, 1, ..., m$ are unknown nonlinear vector fields. The identifier design for (1) is studied under the following basic assumptions:

**Assumption 1** *The vector fields $f_i, i = 0, 1, ..., m$ are continuous on the unit sphere $S = \{x \in \mathbb{R}^n : \|x\| = 1\}$.*

**Assumption 2** *The vector fields $f_i$ are homogeneous in the standard sense with known homogeneity degrees $v_i \in \mathbb{R}$, i.e., $f_i(\lambda x) = \lambda^{v_i} f_i(x), \forall x \in \mathbb{R}^n, \forall \lambda > 0$, where $i = 0, 1, ..., m$.*

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

**Remark 1** *Notice that the generalized concepts of homogeneity [27,29–31] can be considered instead of the standard one. In this case, the condition Assumption 2 becomes $f_i(D(\lambda)x) = \lambda^{v_i} D(\lambda) f_i(x), \forall x \in \mathbb{R}^n, \forall \lambda > 0$, where $D : \mathbb{R}_+ \to \mathbb{R}^{n \times n}$ is a dilation in $\mathbb{R}^n$ (see [31,32] for more details). For $D(\lambda) = \lambda I$ the uniform dilation is recovered. In the finite dimensional case, any generalized homogeneous system is topologically equivalent to a standard homogeneous one [32,33], i.e. Assumption 2 holds after certain transformation of the coordinates.*

**Assumption 3** *The whole state vector of* (1) *assumed to be on-line measured, bounded and sufficiently excited by control inputs (the details are given in Theorems 2 and 3).*

**Assumption 4** *The control inputs are known essentially bounded functions such that $|u_i(t)| \leq U < +\infty$, $\forall t \in \mathbb{R}_+$, $i = 1, \ldots, m$.*

**Remark 2** *Assumptions 3 and 4 are commonly used in system identification (parametric and non-parametric). There are new studies working on incomplete input-output measurement or relaxed excitation concepts [1,34–36]. Since it is the first time, when a DNN is proposed for homogeneous systems, we decided to focus on the most common hypotheses leaving technical relaxations for further researches.*

The aim of this study is to solve the identification problem for (1) under the proposed assumptions 1-4. The solution consists of finding an approximate model for the vector field $f$ and its parameters such that, the error between the states of (1) and the states of the approximate structure are small (in a certain norm). Therefore, the first step is to represent (1) as a valid approximate model taking into account the homogeneity property of $f$. This work proposes a homogeneous DNN structure. The second step consists in designing adaptive laws for the adjustment of weights of this DNN identifier, such that, the error $e := x - \hat{x}$ 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 $\varepsilon^+ = 0$ (i.e. the DNN model may exactly approximate the original system) then the error $e(t)$ must tend to zero as $t \to +\infty$.

**Remark 3** *Homogeneity allows local properties of vector fields to be extended globally, for example, the vector field $f_i : \mathbb{R}^n \to \mathbb{R}^n$ satisfying Assumption 2, is locally Lipschitz continuous on $\mathbb{R}^n \setminus \{0\}$ if and only if it satisfies the Lipschitz condition on the unit sphere, see e.g. [32] and [37]. Similarly, (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 (at the origin) models.*

## 3 Neural networks approximation property for homogeneous systems

The universal approximation property of ANN has been used in several works, which claims that any continuous function can be approximated arbitrarily closely on a compact set using weighted superposition of nonlinear functions such as polynomials, radial basis functions [38] and sigmoidal functions [3,10,12,39]. The following theorem justifies the approximation of continuous functions by sigmoidal functions:

**Theorem 1** *[39] Let $\sigma(\cdot)$ be a non-constant, bounded monotone-increasing continuous function. Let $K$ be a compact subset of $\mathbb{R}^n$ and $f(x_1, \dots, x_n)$ be a real valued continuous function on $K$. Then, for an arbitrary $\varepsilon > 0$, there exists an integer $N$ and real constants $\omega_i$, $c_i$, $b_{ij}$, where $i = 1, \dots, N$ and $j = 1, \dots, n$, such that $F(x) := \sum_{i=1}^{N} \omega_i \sigma \left( \sum_{j=1}^{n} b_{ij} x_j + c_i \right)$ satisfies $\max_{x \in K} |F(x) - f(x)| < \varepsilon$.*

Notice that Theorem 1 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 [40]. This property is used in DNN structures for the representation of dynamic systems.

**Remark 4** *In [7,41] and [42], it is stated that the parameters $c_i$, and $b_{ij}$ can be selected randomly using a uniform distribution. Then, in the aforementioned works, it is shown that the universal approximation property holds by finding only the parameters $\omega_i$.*

All classical DNN identifiers are not homogeneous by construction, so the dilation symmetry of the system (1) cannot be preserved in an approximate model. Since any homogeneous vector field is uniquely defined by its values on the unit sphere and the homogeneity degree, the above theorem implies the following approximation property for the system (1).

**Corollary 1** *Let (1) satisfy the assumptions 1 and 2. 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 $W_i^* \in \mathbb{R}^{n \times N_i}$, $i = 0, 1, \dots, m$ such that:*

$$\|p(x,u)\| \le \varepsilon_0 \|x\|^{\nu_0} + \sum_{i=1}^{m} \varepsilon_i \|x\|^{\nu_i} |u_i|, \quad \forall x \in \mathbb{R}^n, \forall u \in \mathbb{R}^m, \quad (2)$$

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

$$F(x,u) = \|x\|^{\nu_0} \left[ \frac{Ax}{\|x\|} + W_0^* \sigma_0 \left( \frac{x}{\|x\|} \right) \right] + \sum_{i=1}^{m} \|x\|^{\nu_i} W_i^* \sigma_i \left( \frac{x}{\|x\|} \right) u_i. \quad (3)$$

*The elements of the vector functions $\sigma_i : \mathbb{R}^n \to \mathbb{R}_i^N$ are proposed as sigmoidal activation functions, that is:*

$$(\sigma_i(x))_j = \left( 1 + c_{ij} e^{-b_{ij}^\top x} \right)^{-1}, \quad (4)$$

*where $c_{ij} \in \mathbb{R}_+$ and $b_{ij} \in \mathbb{R}^n$ are properly selected parameters with $i = 0, 1, \dots, m$ and $j = 1, \dots, N_i$.*

**PROOF.** Based on the properties given in Assumption 2, the system (1) can be rewritten as $\dot{x} = \|x\|^{\nu_0} f_0(x/\|x\|) + \sum_{i=1}^{m} \|x\|^{\nu_i} f_i(x/\|x\|) u_i$. Notice that, the right-hand side of the system (1) is uniquely identified by its values on the unit sphere.

Consider the functions $\tilde{f}_i : \mathbb{R}^n \to \mathbb{R}^n$, $i = 0, 1, \dots, m$ (associated to the components of the control $u$, except $\tilde{f}_0$), that is $\tilde{f}_i(x) = f_i(x) = f_i(x/\|x\|)$. 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, see Remark 3. Applying the Theorem 1 to each component of the vector $\tilde{f}_i$, $i = 1, 2, \dots m$ and to $\tilde{f}_0(x) - Ax/\|x\|$ on the unit sphere $S$, 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 $\varepsilon_i$. 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) $W_i^*$, $b_{ij}$ and $c_{ij}$. Taking into account the homogeneity of nonlinear functions $f_i, i = 0, \dots, m$, the desired global estimate (2) holds.

It is a well-known fact that usually an ANN guarantees a certain quality of approximation on a compact set only (that happens out this set is often not specified, and the error may become quickly growing) (see, [7,12,39]). Corollary 1 states results on approximation by the proposed homogeneous ANN not just for a particular compact set but quantifies how the approximation error grows globally. Hence, and in view of Remark 4, below we assume that the parameters $b_{ij}$ and $c_{ij}$ are selected randomly with a uniform distribution, and it is needed to find matrices $W_i^*$ in order to complete the identification of the homogeneous model.

**Remark 5** *The activation functions (4) are bounded. Hence, the vectors of the activation functions used for the identification on the unit sphere are bounded in the following sense:*

$$\|\sigma_i(x/\|x\|)\| \le \sqrt{N_i}, \; \forall x \in \mathbb{R}^n. \quad (5)$$

**Remark 6** *Usually, a function $f$ can be approximated by an ANN structure only on a compact set (see Theorem 1). The homogeneous ANN structure (3) gives a global approximation of $f$ (see the formula (2)).*

## 4 Identification of affine homogeneous control systems

### 4.1 The case of known control gains

Let us consider initially the case when the nonlinear vector fields $f_i : \mathbb{R}^n \to \mathbb{R}^n$, $i = 1, ..., m$ associated with the input are known, so *it is only necessary to identify the vector field $f_0$*.

**Remark 7** *In Corollary 1, 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 $f_i, i = 0, ..., m$ as known, leads to the following inequality:*

$$\|d(x/\|x\|)\| \le \varepsilon_0 , \ \forall x \in \mathbb{R}^n, \tag{6}$$

*where $d(x/\|x\|) = f_0(x/\|x\|) - Ax/\|x\| - W_0^* \sigma_0(x/\|x\|)$.*

Notice that $W_0^* \sigma_0(x/\|x\|) := \Sigma_0(x/\|x\|) w_0^*$, where $\Sigma_0(z) = I_n \otimes \sigma_0^\top(z) \in \mathbb{R}^{n \times nN_0}$, $z \in \mathbb{R}^n$, $I_n \in \mathbb{R}^{n \times n}$ is the identity matrix and $w_0^* = \text{vec}\left((W_0^*)^\top\right) \in \mathbb{R}^{nN_0}$. The Frobenius norm of the matrix $\Sigma_0(x/\|x\|)$ has a finite upper-bound (in the matrix space) in view of (5), i.e.

$$\|\Sigma_0(x/\|x\|)\|_F \le \sqrt{nN_0}, \ \forall x \in \mathbb{R}^n. \tag{7}$$

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 (8)-(10).

**Theorem 2** *Let assumptions 1-4 be satisfied and consider the approximation (3) for the homogeneous vector field $f_0$, with modeling error as in Remark 7 and*

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

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

$$\frac{d}{dt}w_0 = -\|x\|^{\nu_0}K\Omega^\top e, \tag{9}$$

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

$$\frac{d}{dt}\Omega = \|x\|^{\nu_0}\left(A\Omega/\|x\| - \Sigma_0(x/\|x\|)\right). \tag{10}$$

*If $K \in \mathbb{R}^{nN_0 \times nN_0}$ is a positive definite symmetric matrix and the control inputs $u$ in (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 holds $\int_t^{t+\ell_\omega} \Omega^\top(s)\Omega(s)ds \ge \vartheta_\omega I_{nN_0}$, $\forall t \in \mathbb{R}_+$, for some $\ell_\omega > 0$ and $\vartheta_\omega > 0$. Then, there exist two class-$\mathcal{K}$ functions $\rho_1$ and $\rho_2$ such that $\limsup_{t \to \infty} \|e(t)\| \le \rho_1(\varepsilon_0)$, and*

$$\limsup_{t \to \infty} \|w_0(t) - w_0^*(t)\| \le \rho_2(\varepsilon_0), \tag{11}$$

*where $\varepsilon_0$ is given by (6).*

**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:

$$\delta = e + \Omega\tilde{\omega}. \tag{12}$$

The dynamics of (12) satisfies:

$$\dot{\delta} = \dot{e} + \frac{d\Omega}{dt}\tilde{\omega} + \Omega\frac{d\tilde{\omega}}{dt}. \tag{13}$$

In (13), $\dot{e}$ corresponds to:

$$\dot{e} = \|x\|^{\nu_0}\left[\frac{Ae}{\|x\|} + \Sigma_0\left(\frac{x}{\|x\|}\right)\tilde{\omega} - \Omega K\Omega^\top e + d\left(\frac{x}{\|x\|}\right)\right]. \tag{14}$$

The time derivative of $\tilde{\omega}$ satisfies $\frac{d\tilde{\omega}}{dt} = -\frac{dw_0}{dt}$. Then, the substitution of (9), (10) and (14) on (13) yields:

$$\dot{\delta} = \|x\|^{\nu_0}\left[A\frac{e}{\|x\|} + \Sigma_0\left(\frac{x}{\|x\|}\right)\tilde{\omega} - \Omega K\Omega^\top e + d\left(\frac{x}{\|x\|}\right)\right] \\ + \|x\|^{\nu_0}\left(A\frac{\Omega}{\|x\|} - \Sigma_0\left(\frac{x}{\|x\|}\right)\right)\tilde{\omega} + \|x\|^{\nu_0}\Omega K\Omega^\top e. \tag{15}$$

According to the definition of $\delta$ in (12), (15) is equivalent to $\dot{\delta} = \|x\|^{\nu_0}((A/\|x\|)\delta + d(x/\|x\|))$, since $A$ is Hurwitz, $\|x\| \ne 0$ and the modelling error $d(\cdot)$ corresponds to a bounded additive input (6) in $\dot{\delta}$, we can conclude that the auxiliary variable $\delta$ is also bounded, i.e., $\limsup_{t \to \infty} \|\delta(t)\| \le \rho_0(\varepsilon_0)$, where $\rho_0$ is a class-$\mathcal{K}$ function (see Lemma 1 in the Appendix). By the same result and considering (7), as well as (10), it is straightforward to observe that the variable $\Omega$ is also bounded, i.e. $\limsup_{t \to \infty} \|\Omega(t)\|_F \le \bar{\Omega}$, where $\bar{\Omega}$ is depends on $\sqrt{nN_0}$ (see (7)), $x^+$ and $x^-$. Using the selected learning law (9) and (12), one obtains:

$$\frac{d\tilde{\omega}}{dt} = -\|x\|^{\nu_0}K\Omega^\top(\Omega\tilde{\omega} - \delta). \tag{16}$$

Due to boundedness $\|x\|$ and $\Omega$, the system (16) is input-to-state stable with respect to the input $\delta$ (see Corollary 2 in the Appendix), i.e. the inequality (11) holds. An estimate for the convergence quality of the identification error can be obtained using (12). Since $e = \delta - \Omega\tilde{\omega}$, one may notice that:

$$\|e\| \le \|\delta\| + \|\Omega\tilde{\omega}\|, \ \forall t \in R_+. \tag{17}$$

Using the norm relation in (17) (see e.g. [43]) and the fact that bound of $\tilde{\omega}$ depends on the upper bound for $\delta$ (see Lemma 1 in the Appendix), one gets $\limsup_{t \to \infty} \|\Omega(t)\tilde{\omega}(t)\| \le \limsup_{t \to \infty} \|\tilde{\omega}(t)\|\bar{\Omega} \le \bar{\Omega}\rho_2(\varepsilon_0)$, where $\rho_2$ is a class-$\mathcal{K}$ function from (11). The following relation for (17) takes place, $\limsup_{t \to \infty} \|e(t)\| \le \rho_1(\varepsilon_0) := \rho_0(\varepsilon_0) + \bar{\Omega}\rho_2(\varepsilon_0)$. The proof is complete.

Notice that the DNN identifier (8)-(10) is not regular [8,16] because it has a direct injection of the identification error $e$. Notice also that in many cases the PE condition can be fulfilled by a proper selection of a control input $u$ (see e.g. [44] for more details about this).

### 4.2 The case of unknown control gains

In this section, we design a DNN identifier for the considered system assuming that functions $f_i$ are unknown but admit the representation (3). Similarly to the previous case, we introduce the vectors $w_i^*$ and the matrix-valued function $\Sigma_i$ such that $W_i^* \sigma_i(x/\|x\|) := \Sigma_i(x/\|x\|) w_i^*$. The matrices $\Sigma_i(x/\|x\|)$ have a finite upper bound (in the matrix space) in view of (5), i.e. $\|\Sigma_i(x/\|x\|)\|_F \le \sqrt{nN_i}, \ \forall x \in \mathbb{R}^n$.

**Theorem 3** *Let assumptions 1-3 be satisfied and the control input $u$ be selected as $u_i(t) = \frac{\tilde{u}_i(t)}{\|x(t)\|^{v_i - v_0}}, \quad i = 1,\ldots,m,$ where $\tilde{u}_i : \mathbb{R}_+ \to \mathbb{R}$ are continuous bounded uniformly bounded functions, i.e. that $|\tilde{u}_i(t)| \le \tilde{U}, \forall t \in \mathbb{R}_+$ for some number $\tilde{U} > 0$.*

*Consider the system (1) which can be represented in the form (3) with an estimation error given in (2). Define*

$$\frac{d\hat{x}}{dt} = \|x\|^{v_0-1} A \hat{x} + \|x\|^{v_0} \left[ \sum_{i=0}^{m} \Sigma_i \left( \frac{x}{\|x\|} \right) w_i \tilde{u}_i + \Omega_i K_i \Omega_i^\top e \right],$$

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

$$\frac{d}{dt} w_i = -\|x\|^{v_0} K_i \Omega_i^\top e \tag{18}$$

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

$$\frac{d}{dt} \Omega_i = \|x\|^{v_0-1} A \Omega_i - \|x\|^{v_0} \tilde{u}_i \Sigma_i(x/\|x\|), \tag{19}$$

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

$$\int_t^{t+\ell_W} G^\top(s) G(s) ds \ge \vartheta_W I_{n \sum_{i=0}^m N_i},$$

*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 $\rho_1$ and $\rho_2$, such that $\limsup_{t \to \infty} \|e(t)\| \le \rho_1(\varepsilon^+)$ and*

$$\limsup_{t \to \infty} \|w_i(t) - w_i^*(t)\| \le \rho_2(\varepsilon^+), \tag{20}$$

*where $\varepsilon^+ = \max_{i=0,\ldots,m} \{\varepsilon_i\}$ and $\varepsilon_i$ are given by (2).*

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

$$\delta = e + \sum_{i=0}^{m} \Omega_i \tilde{\omega}_i. \tag{21}$$

Hence, the dynamics of (21) is:

$$\dot{\delta} = \dot{e} + \sum_{i=0}^{m} \left[ \frac{d\Omega_i}{dt} \tilde{\omega}_i + \Omega_i \frac{d\tilde{\omega}_i}{dt} \right], \tag{22}$$

where $\dot{e}$ is given by:

$$\dot{e} = \|x\|^{v_0-1} A e + p(x,u)$$
$$+ \|x\|^{v_0} \sum_{i=0}^{m} \left[ \Sigma_i \left( \frac{x}{\|x\|} \right) \tilde{\omega}_i \tilde{u}_i - \Omega_i K_i \Omega_i^\top e \right], \tag{23}$$

the function $p(x,u)$ is defined in (2). The time derivative of $\tilde{\omega}_i$ satisfies $\frac{d\tilde{\omega}_i}{dt} = -\frac{dw_i}{dt}$. Then, the substitution of (18), (19) and (23) on (22) yields to:

$$\dot{\delta} = \|x\|^{v_0-1} A e + \|x\|^{v_0} \sum_{i=0}^{m} \Omega_i K_i \Omega_i^\top e$$
$$+ \|x\|^{v_0} \sum_{i=0}^{m} \left[ \Sigma_i \left( \frac{x}{\|x\|} \right) \tilde{\omega}_i \tilde{u}_i - \Omega_i K_i \Omega_i^\top e \right] + p(x,u) \tag{24}$$
$$+ \sum_{i=0}^{m} \left( \|x\|^{v_0-1} A \Omega_i - \|x\|^{v_0} \tilde{u}_i \Sigma_i \left( \frac{x}{\|x\|} \right) \right) \tilde{\omega}_i.$$

According to (21), (24) is equivalent to $\dot{\delta} = \|x\|^{v_0-1} A \delta + p(x,u)$. Taking into account the identity $u_i(t) = \tilde{u}(t)/\|x\|^{v_i-v_0}$ from (2) we derive $\|p(x(t), u(t))\| \le \|x\|^{v_0} (\varepsilon_0 + \tilde{U} \sum_{i=1}^m \varepsilon_i)$. Since $A$ is Hurwitz and $x$ is uniformly bounded from below and from above, then using Lemma 1 we have $\limsup_{t \to \infty} \|\delta(t)\| \le \rho_0(\varepsilon^+)$, for some class-$\mathcal{K}$ function $\rho_0$. Similarly, for (19) we conclude $\limsup_{t \to \infty} \|\Omega_i(t)\| \le \bar{\Omega}_i$, where the number $\bar{\Omega}_i$ depends on $\sqrt{nN_i}$, $\tilde{U}$, $x^-$ and $x^+$ (see Lemma 1 in the Appendix). Using (18) and (21), we obtain:

$$\frac{d\tilde{\omega}_i(t)}{dt} = -\|x\|^{v_0} K_i \Omega_i^\top \left( \sum_{j=0}^{m} \Omega_j \tilde{\omega}_j - \delta \right).$$

or, equivalently, $\frac{dW}{dt} = -\|x\|^{v_0} KG^\top (GW - \delta)$, where $W(t) = \left[ \tilde{\omega}_0^\top(t), \tilde{\omega}_1^\top(t), \ldots, \tilde{\omega}_m^\top(t) \right]^\top$, $K = \text{diag}\{K_0, K_1, \ldots, K_m\}$. Due to boundedness of $\Omega_i$, from Corollary 2 we derive (20). Finally, using (21) we finish the proof:

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

## 5 Numerical Results

To show the performance of the proposed algorithm a three-tank system as the depicted in Figure 1 is used. The same figure includes the variables and parameters considered in the dynamic model whose dynamics is described by the following homogeneous nonlinear system [22]:

$$\dot{x}_1 = S_{tank}^{-1} \left[ -\theta_1 \lceil x_1 - x_3 \rfloor^{0.5} + u_1 \right],$$
$$\dot{x}_2 = S_{tank}^{-1} \left[ \theta_3 \lceil x_3 - x_2 \rfloor^{0.5} - \theta_2 \lceil x_2 \rfloor^{0.5} + u_2 \right],$$
$$\dot{x}_3 = S_{tank}^{-1} \left[ \theta_1 \lceil x_1 - x_3 \rfloor^{0.5} - \theta_3 \lceil x_3 - x_2 \rfloor^{0.5} \right],$$

where $x_1$ [m], $x_2$ [m] and $x_3$ [m] 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 $\theta_1$, $\theta_2$ and $\theta_3$, [$m^{3/2}/s^2$] are coefficients related with the outflow rate according to Torricelli's rule.
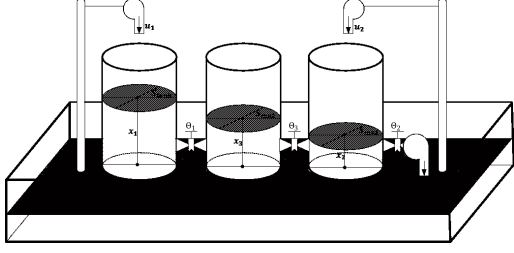
5

Fig. 1. Three-tank system

Tank systems have been used to test numerous control and identification algorithms [22,45,46]. For simulation we considered the following parameters $S_{tank} = 1$, $\theta_1 = 3$, $\theta_3 = 2$, $\theta_2 = 1$, $x_0 = [3, 1, 2]^\top$ as the initial condition. 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:

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

The parameters associated with the activation functions given by (4) were selected randomly with a uniform distribution (See Remark 4). To adjust these parameters, recursive procedures could be used to obtain a better performance [42,47,48]. The activation functions were selected with $N_i = 3$. The constant parameters $c_{0,1} = 5$, $c_{0,2} = 10$ and $c_{0,3} = 20$, the constant vectors $b_{0,1} = [0.01, 0.02, 0.03]^\top$, $b_{0,2} = [0.01, 0.04, 0.01]^\top$, $b_{0,3} = [0.04, 0.01, 0.06]^\top$. Equal constant parameters $c_{i,j}$ and $c_{0,j}$ are chosen ($c_{i,j} = c_{0,j}$) as well as in the case of the vectors $b_{i,j} = b_{0,j}$. The initial conditions for the adjustment laws are selected as: $w_0(0) = [1,2,1,1,3,2,1,2,3]^\top$, $w_1(0) = [1,3,1,5,3,2,1,2,3]^\top$ and $w_2(0) = [1,4,1,2,1,6,1,3,3]^\top$, the gain matrices $K_0 = I_3 \otimes \tilde{K}_0$ and $K_1 = K_2 = I_3 \otimes \tilde{K}_1$, with:

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

which were obtained randomly, satisfying the positiveness condition needed for the activation functions. The performance of the designed algorithm was compared with a classical DNN identifier (see [8], 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*. The weights of the proposed identifier followed the laws (18) and the classical DNN weights followed the laws presented in [8], Chapter 2. 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 linear inequality for the classical DNN identifier structure is used (together with a numerical optimization algorithm) to get some of the parameters used in the learning law as reported in [8]. The adjustment gains were selected manually using the simulations results, fixed with $K = [55, 65.2]$. The initial conditions are the same for both identifier algorithms $\hat{x}(0) = [5, 2, 3]^\top$.

Figure 2-a) depicts how the first estimate state of the identifier for homogeneous systems converges faster (before 0.01 seconds) to the state of the system (solid line, color black) than the classical DNN identifier. The estimation of the classical algorithm contains oscillations with an estimated amplitude of 2 [m] before 0.5 seconds. The amplitude of these oscillations is unacceptable considering the nature of the three-tank system. The lack of oscillations and the faster convergence in the state evolution of the homogeneous identifier is clearly an advantage.
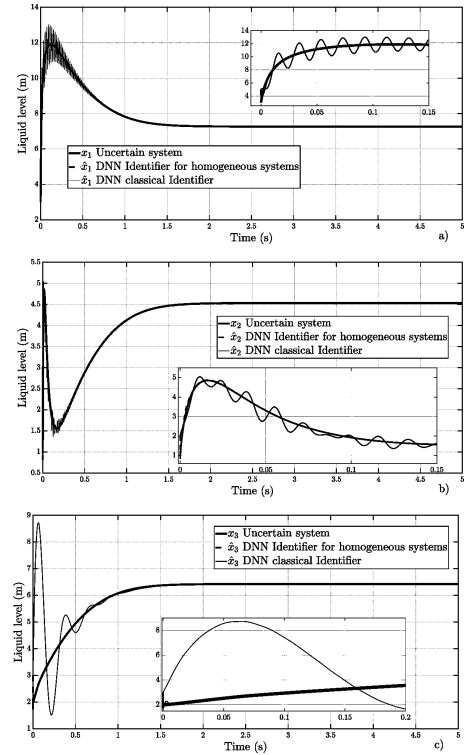


Fig. 2. Identification result for the states of the three tank system.

Figure 2-b) 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 classical series-parallel DNN structure presents bigger oscillations and has a slower convergence than the homogeneous algorithm. In Figure 2-c), 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 1.2 seconds). In addition, the estimation with the classical algorithm presents oscillations. Although the oscillations are of

a smaller frequency than the other two states, the amplitude of such oscillations is approximately of 3.5 [m] before 0.8 seconds. In addition, the enclosed figure confirms the equal initial conditions for both identifiers.

In Figure 3-a), 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 unknown model (solid line, color blue) is depicted. The identification error with the classical DNN identifier converges around 1.5 seconds versus 0.02 seconds of the identification error when the DNN identifier is devoted to homogeneous systems. The comparison sums up the performance of the states obtained with the classical and homogeneous identifiers. The obtained identification error norm with the classical approach presents oscillations, which are not detected in the homogeneous version.
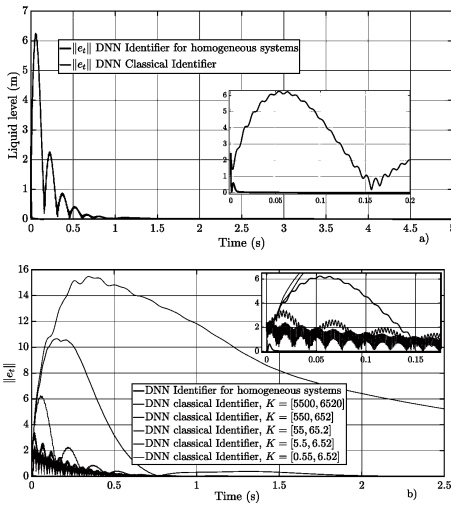


Fig. 3. Comparison of the identification error norms for the homogeneous and classical DNN identifiers.

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. In addition, Figure 3-b) depicts the result of a different gain test for the classical algorithm. In order to increase the convergence rate of the classical DNN, the gain $K$ was increased, the latter implied the appearance of the high-frequency oscillations and it did not improve the convergence rate compared with the identifier for homogeneous systems (blue line and black line). On the other hand, the decrease of the gain $K$ decelerates the convergence (purple line and orange line). This result shows the faster convergence of the homogeneous identifier.

Intending to make a comparison with conventional approaches, we carried out a statistical analysis using the classical DNN identifier and the proposed homogeneous identifier. The integral square mean error calculated after

2.5 seconds is 0.001469 [m$^2$] for the homogeneous and 0.3589 [m$^2$] for the non-homogeneous identifiers. In addition, Table 1 shows the average and standard deviation of the final value of the accumulated total error using different initial conditions and using different state evolution velocities, for these results we performed ten tests.

Table 1
Statistical results comparing classical and homogeneous DNN.

| Identifier | Average | Standard Deviation |
|---|---|---|
| Classical DNN with different initial weights | 1.6667 | 0.9598 |
| Homogeneous DNN with different initial weights | 0.0052 | 0.0004 |
| Classical DNN with different states evolution | 1.19 | 0.3606 |
| Homogeneous DNN with different states evolution | 0.0060 | 0.0074 |

## 6  Conclusions

The design of an identification algorithm for standard homogeneous control systems was developed in this paper 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. Numerical simulations are presented using a three thank homogeneous model to demonstrate the performance of the identification algorithm. As a future extension to this work, 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 2, as well as the implementation in an experimental platform. Relaxation of Assumption 3 is another promising direction of research.

## A  Proof of boundedness for the auxiliary variables

**Lemma 1** *Consider the dynamical system $\dot{r} = \varphi_1(t)Ar + \varphi_2(t)$, 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:*

$$\varphi_1(t) \geq \beta_1 > 0, \quad \forall t \in \mathbb{R}_+, \tag{A.1}$$

*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 the system satisfy:*

$$\limsup_{t \to \infty} \|r(t)\| \leq \rho_r(\varphi_2^+). \tag{A.2}$$

**PROOF.** Consider the Lyapunov function candidate $V_a(r) = r^\top Pr$. The time derivative is $\dot{V}_a = \varphi_1 r^\top (A^\top P + PA) r + 2\varphi_2^\top Pr$, where $P \in \mathbb{R}^{n \times n}$, $P = P^\top > 0$, such that:

$$A^\top P + PA \leq -\gamma P, \ \gamma \in \mathbb{R}_+, \tag{A.3}$$

Using the $\Lambda$-inequality we derive $2\varphi_2^\top Pr \leq \|\varphi_2\|_\Lambda^2 + \|Pr\|_{\Lambda^{-1}}^2$, for any $\Lambda \in \mathbb{R}^{n \times n}$, $\Lambda = \Lambda^\top$, $\Lambda > 0$. By taking $\Lambda = 2\beta_1^{-1}\gamma^{-1}P$, considering (A.3) and the condition (A.1). Then, $\dot{V}_a \leq -0.5\beta_1\gamma V_a(r) + 2\beta_1^{-1}\gamma^{-1}\lambda_{\max}\{P\}\|\varphi_2(t)\|^2$ holds. Hence, we derive $\limsup_{t \to +\infty} V_a(t) \leq \frac{2\beta_1^{-1}\gamma^{-1}\lambda_{\max}\{P\}(\varphi_2^+)^2}{0.5\beta_1\gamma}$, and (A.2) is fulfilled.

## B    Persistent excitation condition

The following result is a Corollary of [44, Lemma 1].

**Corollary 2** *Consider the time-varying dynamical system:*

$$\dot{z}(t) = -\gamma_z(t)K_z G_z^\top(t)G_z(t)z(t) + v(t),\ t \in \mathbb{R}_+ \qquad \text{(B.1)}$$

*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*

$$\int_t^{t+\ell} G_z^\top(s)G_z(s)ds \geq \vartheta I_k, \quad \forall t \in \mathbb{R}_+ \qquad \text{(B.2)}$$

*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 (B.1) 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: $\frac{d}{dt}\tilde{z}(t) = -\gamma_z(t)K_z^{1/2}G_z^\top(s)G_z(s)K_z^{1/2}\tilde{z}(t) + K_z^{-1/2}v(t)$. Then, $\frac{d}{dt}\tilde{z}(t) = -\tilde{G}^\top(s)\tilde{G}(s)\tilde{z}(t) + \tilde{v}(t)$, 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 $\int_t^{t+\ell} \tilde{G}^\top(s)\tilde{G}(s)ds = K_z^{1/2}\left(\int_t^{t+\ell} \gamma_z(s)G_z^\top(s)G_z(s)ds\right)K_z^{1/2}$. Then, $\int_t^{t+\ell} \tilde{G}^\top(s)\tilde{G}(s)ds \geq \gamma^- K_z^{1/2}\left(\int_t^{t+\ell} G_z^\top(s)G_z(s)ds\right)K_z^{1/2}$ and $\int_t^{t+\ell} \tilde{G}^\top(s)\tilde{G}(s)ds \geq \gamma^- K_z^{1/2}(\vartheta I_k)K_z^{1/2} = \gamma^-\vartheta K_z \geq \gamma^-\vartheta\lambda_{\min}(K_z)I_k$, $\forall t \in \mathbb{R}_+$. Applying [44, Lemma 1] we complete the proof.

## References

[1]  L. Ljung, *System Identification Theory for the User*, ser. Prentice Hall Information and System Sciences Series, J. Gwyn, Ed.  Prentice Hall PTR, 1999.

[2]  O. Nelles, *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*, ., Ed.  Springer Science & Business Media, 2013.

[3]  J. A. Farrell and M. M. Polycarpou, *Adaptive approximation based control: unifying neural, fuzzy and traditional adaptive approximation approaches*.  John Wiley & Sons, 2006, vol. 48.

[4]  A. Simorgh, A. Razminia, and V. I. Shiryaev, "System identification and control design of a nonlinear continuously stirred tank reactor," *Mathematics and Computers in Simulation*, vol. 173, pp. 16 – 31, 2020.

[5]  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.

[6]  S. Haykin, *Neural networks: a comprehensive foundation*.  Prentice Hall PTR, 1994.

[7]  F. Lewis, S. Jagannathan, and A. Yesildirak, *Neural network control of robot manipulators and non-linear systems*, ., Ed.  CRC Press, 1998.

[8]  A. S. Poznyak, E. N. Sanchez, and W. Yu, *Differential neural networks for robust nonlinear control: identification, state estimation and trajectory tracking*, ., Ed.  World Scientific, 2001.

[9]  D. Chen, S. Li, Q. Wu, and L. Liao, "Simultaneous identification, tracking control and disturbance rejection of uncertain nonlinear dynamics systems: A unified neural approach," *Neurocomputing*, vol. 381, pp. 282 – 297, 2020.

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

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

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

[13]  S. Hubbert, "Radial basis function interpolation on the sphere," Ph.D. dissertation, University of London, 2002.

[14]  H. V. H. Ayala and L. dos Santos Coelho, "Cascaded evolutionary algorithm for nonlinear system identification based on correlation functions and radial basis functions neural networks," *Mechanical Systems and Signal Processing*, vol. 68-69, pp. 378 – 393, 2016.

[15]  H. V. H. Ayala, D. Habineza, M. Rakotondrabe, and L. dos Santos Coelho, "Nonlinear black-box system identification through coevolutionary algorithms and radial basis function artificial neural networks," *Applied Soft Computing*, vol. 87, p. 105990, 2020.

[16]  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.

[17]  R. Kumar and S. Srivastava, "Externally Recurrent Neural Network based identification of dynamic systems using Lyapunov stability analysis," *ISA Transactions*, vol. 98, pp. 292 – 308, 2020.

[18]  H. Hermes, "Nilpotent approximations of control systems and distributions," *SIAM journal on control and optimization*, vol. 24, no. 4, pp. 731–736, 1986.

[19]  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]  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.

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

[22]  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.

[23]  J.-B. Pomet and C. Samson, "Time-varying exponential stabilization of nonholonomic systems in power form," Inria, Research Report RR-2126, 1993.

[24]  H. Hermes, "Homogeneous feedback controls for homogeneous systems," *Systems & Control Letters*, vol. 24, no. 1, pp. 7 – 11, 1995.

[25]  L. Rosier, "Homogeneous Lyapunov function for homogeneous continuous vector field," *Systems & Control Letters*, vol. 19, pp. 467–473, 1992.

[26]  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.

[27]  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.

[28]  M. Ballesteros, A. Polyakov, D. Efimov, I. Chairez, and A. Poznyak, "Differential Neural Network Identification for Homogeneous Dynamical Systems," *IFAC-PapersOnLine*, vol. 52, no. 16, pp. 233–238, 2019.

[29] 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.

[30] V. V. Khomenuk, "(in russian)," *Izvestiya Vuzov, Matematika*, vol. 3, no. 22, pp. 157–164, 1961.

[31] M. Kawski, "Geometric homogeneity and stabilization," in *Nonlinear Control Systems Design 1995*. Elsevier, 1995, pp. 147–152.

[32] A. Polyakov, "Sliding mode control design using canonical homogeneous norm," *International Journal of Robust and Nonlinear Control*, vol. 0, no. 0, 2018.

[33] L. Grüne, "Homogeneous state feedback stabilization of homogenous systems," *SIAM Journal on Control and Optimization*, vol. 38, no. 4, pp. 1288–1308, 2000.

[34] K. S. Narendra and A. M. Annaswamy, "Persistent excitation and robust adaptive systems," in *Tech. Report# 8302*. Center for Systems Science. Yale University New Haven, Connecticut, 1983.

[35] Y. Song, K. Zhao, and M. Krstic, "Adaptive control with exponential regulation in the absence of persistent excitation," *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2589–2596, 2016.

[36] J. Wang, D. Efimov, and A. Bobtsov, "On robust parameter estimation in finite-time without persistence of excitation," *IEEE Transactions on Automatic Control*, 2019.

[37] 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.

[38] F. Girosi and T. Poggio, "Networks and the best approximation property," *Biological cybernetics*, vol. 63, no. 3, pp. 169–176, 1990.

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

[40] N. E. Cotter, "The Stone-Weierstrass theorem and its application to neural networks," *IEEE Transactions on Neural Networks*, vol. 1, no. 4, pp. 290–295, 1990.

[41] 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.

[42] L. Zhang and P. N. Suganthan, "A comprehensive evaluation of random vector functional link networks," *Information sciences*, vol. 367, pp. 1094–1105, 2016.

[43] D. S. Bernstein, *Scalar, Vector, and Matrix Mathematics: Theory, Facts, and Formulas-Revised and Expanded Edition*. Princeton university press, 2018.

[44] 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.

[45] 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.

[46] G. Giordano and J. Sjöberg, "Black- and white-box approaches for cascaded tanks benchmark system identification," *Mechanical Systems and Signal Processing*, vol. 108, pp. 387 – 397, 2018.

[47] Y.-H. Pao, G.-H. Park, and D. J. Sobajic, "Learning and generalization characteristics of the random vector functional-link net," *Neurocomputing*, vol. 6, no. 2, pp. 163–180, 1994.

[48] D. Husmeier, "Random vector functional link (RVFL) networks," in *Neural Networks for Conditional Probability Estimation*. Springer, 1999, pp. 87–97.